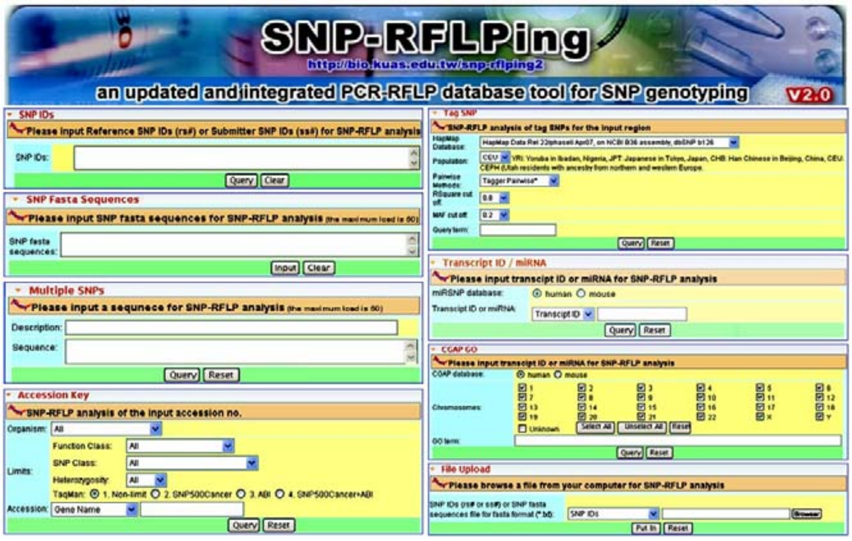
SNP-RFLPing 2

an updated and integrated PCR-RFLP tool for SNP genotyping

User Manual


SNP-RFLPing
http://bio.kuas.edu.tw/snp-rflping2
an updated and integrated PCR-RFLP database tool for SNP genotyping
V2.0
SNP IDs
Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis
SNP IDs:
Query
Clear
SNP Fasta Sequences
Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 60)
SNP fasta sequences:
Input
Clear
Multiple SNPs
Please input a sequnece for SNP-RFLP analysis (the maximum load is 60)
Description:
Sequence:
Query
Reset
Accession Key
SNP-RFLP analysis of the input accession no.
Query
Reset
Tag SNP
SNP-RFLP analysis of tag SNPs for the input region
Query term:
Query
Reset
Transcript ID / miRNA
Please input transcipt ID or miRNA for SNP-RFLP analysis
Query
Reset
CGAP GO
Please input transcipt ID or miRNA for SNP-RFLP analysis
Select All
Unselect All
Reset
GO term:
Query
Reset
File Upload
Please browse a file from your computer for SNP-RFLP analysis
Browse
Put In
Reset

Organism: All

Limits: Function Class: All

SNP Class: All

Heterozygosity: All

TaqMan: ● 1. Non-limit ○ 2. SNP500Cancer ○ 3. ABI ○ 4. SNP500Cancer+ABI

Accession: Gene Name

HapMap Database: HapMap Data Rel 22/phaseII Apr07, on NCBI B36 assembly, dbSNP b126

Population: CEU YRI: Yoruba in Ibadan, Nigeria, JPT: Japanese in Tokyo, Japan, CHB: Han Chinese in Beijing, China, CEU: CEPH (Utah residents with ancestry from northern and western Europe.

Pairwise Methods: Tagger Pairwise*

RSquare cut off: 0.8

MAF cut off: 0.2

miRSNP database: ● human ○ mouse

Transcipt ID or miRNA: Transcipt ID

CGAP database: ● human ○ mouse

Chromosomes: ☑ 1 ☑ 2 ☑ 3 ☑ 4 ☑ 5 ☑ 6
☑ 7 ☑ 8 ☑ 9 ☑ 10 ☑ 11 ☑ 12
☑ 13 ☑ 14 ☑ 15 ☑ 16 ☑ 17 ☑ 18
☑ 19 ☑ 20 ☑ 21 ☑ 22 ☑ X ☑ Y
☐ Unknown

SNP IDs (rs# or ss#) or SNP fasta sequences file for fasta format (*.txt): SNP IDs

Contact:
changhw@kmu.edu.tw
yuhuei.cheng@gmail.com
chuang@isu.edu.tw
chyang@cc.kuas.edu.tw

Updated: 2009/11/03

Table of content

1. Introduction p.2
1.1 Purpose p.2
1.2 System workflow p.2
1.3 Function overview p.3
2. Function: SNP ID for input p.4
2.1 Overview p.4
2.2 Example: A complete operation using SNP ID to analyze the SNP-RFLPing p.4
2.2.1 Example (1): SNP IDs input p.4
2.2.2 Example (2): Query SNPs information p.10
2.2.3 Example (3): SNP-RFLP analysis p.11
2.2.4 Example (4): Natural primers information acquiring p.12
2.2.5 Example (5): Available restriction enzymes request p.13
2.2.6 Example (6): Further restriction enzyme information description p.14
2.2.7 Example (7): Mutagenic primer information acquiring p.17
2.2.8 Example (8): TaqMan information providing p.18
3. Function: SNP in fasta sequence for input p.20
3.1 Overview p.20
3.2 Example: Use SNP in fasta sequence format to analyze p.20
4. Function: Multiple SNPs within one sequence for input p.29
4.1 Overview p.29
4.2 Example: Multiple SNPs within one sequence input p.29
5. Function: GenBank accession no. for input p.33
5.1 Overview p.33
5.2 Example: Use HUGO gene name 'TP53' to analyze p.33
6. Function: TagSNP from HapMap p.35
6.1 Overview p.35
6.2 Example: Use HUGO gene name 'BRCA2' to analyze p.35
7. Function: Transcript ID/miRNA p.37
7.1 Overview p.37
7.2 Example: Use transcript ID and miRNA number to analyze p.37
8. Function: Gene Ontology-based annotation for SNPs p.40
8.1 Overview p.40
8.2 Example: Query Gene Ontology vocabulary term p.40
9. Function: File upload for input p.43
9.1 Overview p.43
9.2 Example: Upload interface for SNP IDs and SNP fasta sequences p.43
10. Appendix p.44
10.1 Definitions p.44
10.2 Related links p.44

1. Introduction

1.1 Background

PCR-restriction fragment length polymorphism (RFLP) assay is a cost-effective method for SNP genotyping and mutation detection, but the manual mining of restriction enzymes is challenging and cumbersome. Three years after we constructed SNP-RFLPing, a freely accessible analysis tool for restriction enzyme mining of SNPs, significant improvements over the 2006 version have been made in the latest version, SNP-RFLPing 2. The primary aim of SNP-RFLPing 2 is to provide a comprehensive PCR-RFLP information for SNPs in multiple fields, such as SNPs for multiple species, multiple kinds (di-, tri, tetra- and indel), gene-centric search, HapMap tagSNPs, gene ontology-based search, miRNAs, and SNP500Cancer. The RFLP restriction enzymes and their corresponding PCR primers of natural and mutagenic types for each SNP are simultaneously analyzed. All the prices of the RFLP restriction enzymes are provided for selection. Furthermore, the previously encountered updating problems for most SNP related databases are resolved by an online retrieval system. The user interfaces for the functional SNP analyses have been substantially improved and integrated. SNP-RFLPing 2 offers a new, user-friendly interface for RFLP genotyping in association studies and is freely available at http://bio.kuas.edu.tw/snp-rflping2/rflpUI.jsp.

1.2 System workflow

The workflow of SNP-RFLPing 2 was illustrated in Figure 1 (see next page). The system was divided into nine modules, including: 1) input module, 2) data retrieve module, 3) sequence process module, 4) SNP-RFLP module, 5) data query module, 6) remote database module, 7) primer design module, 8) output module, and 9) SNP-RFLP database module according to MVC (Model-View-Controller) architecture. Input and Output modules are served with the graphical user interface (GUI) to operate and analyze.

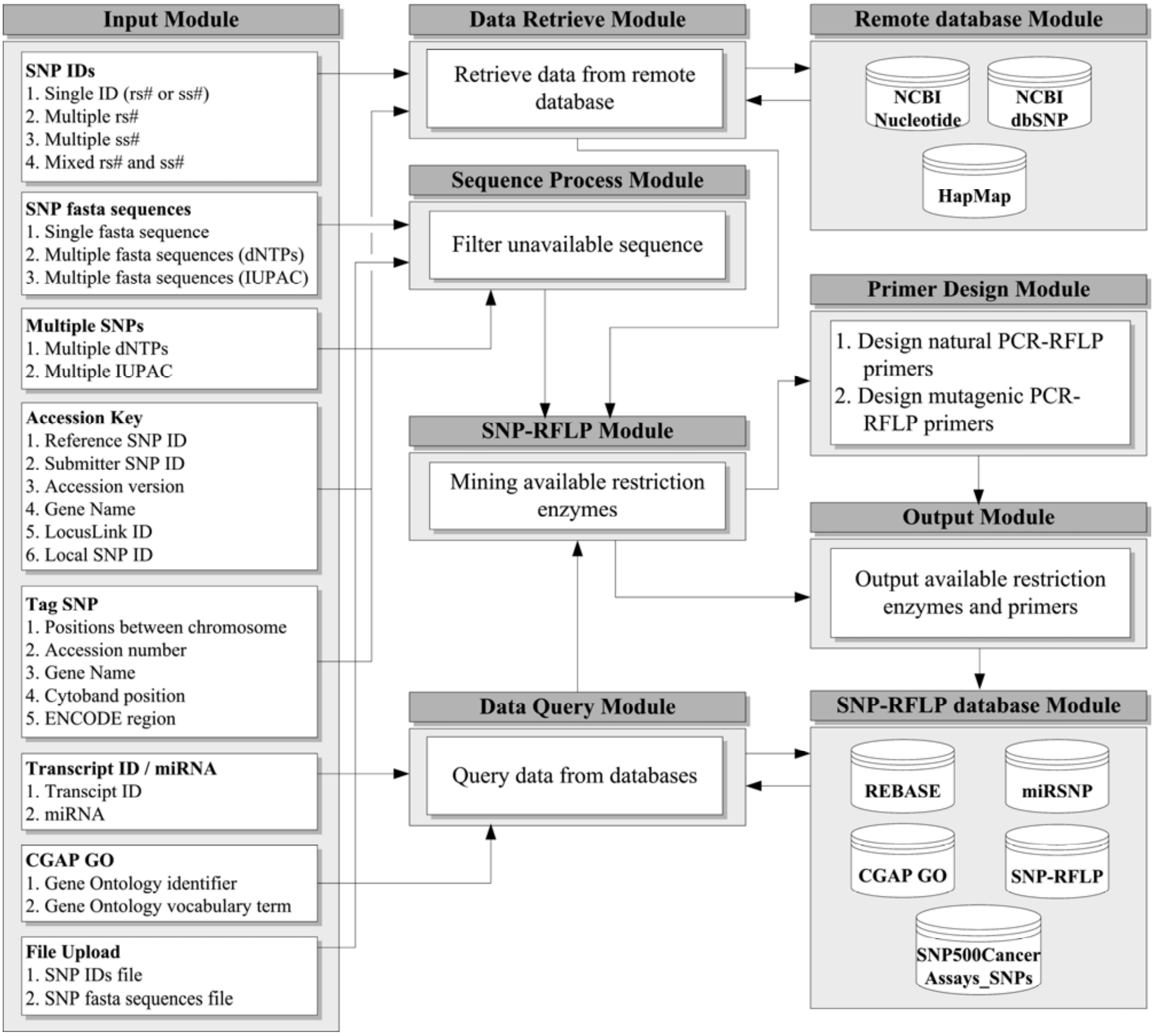

Input Module
SNP IDs
1. Single ID (rs# or ss#)
2. Multiple rs#
3. Multiple ss#
4. Mixed rs# and ss#
SNP fasta sequences
1. Single fasta sequence
2. Multiple fasta sequences (dNTPs)
3. Multiple fasta sequences (IUPAC)
Multiple SNPs
1. Multiple dNTPs
2. Multiple IUPAC
Accession Key
1. Reference SNP ID
2. Submitter SNP ID
3. Accession version
4. Gene Name
5. LocusLink ID
6. Local SNP ID
Tag SNP
1. Positions between chromosome
2. Accession number
3. Gene Name
4. Cytoband position
5. ENCODE region
Transcript ID / miRNA
1. Transcipt ID
2. miRNA
CGAP GO
1. Gene Ontology identifier
2. Gene Ontology vocabulary term
File Upload
1. SNP IDs file
2. SNP fasta sequences file
Data Retrieve Module
Retrieve data from remote database
Sequence Process Module
Filter unavailable sequence
SNP-RFLP Module
Mining available restriction enzymes
Data Query Module
Query data from databases
Remote database Module
NCBI Nucleotide
NCBI dbSNP
HapMap
Primer Design Module
1. Design natural PCR-RFLP primers
2. Design mutagenic PCR-RFLP primers
Output Module
Output available restriction enzymes and primers
SNP-RFLP database Module
REBASE
miRSNP
CGAP GO
SNP-RFLP
SNP500Cancer Assays_SNPs

Fig. 1-1. System structure and flowchart for SNP-RFLPing 2.

1.3 Function overview

SNP-RFLPing 2 provides *eight major functions* to assist and analyze the RFLP restriction enzymes and its corresponding primers for SNPs in all SNP-related fields, including:

(1) SNP ID input

(2) SNP in fasta sequence input

(3) Multiple SNPs within one sequence

(4) GenBank accession

(5) TagSNP from HapMap

(6) Transcript ID/miRNA

(7) Gene Ontology-based annotation for SNPs

(8) File upload

2. Function: SNP ID input

2.1 Overview

SNP ID (rs# and ss#) is acceptable to query the SNP-RFLP information. Single ID, multiple rs#, multiple ss#, and mixed rs# and ss# are available for SNP ID input. For multiple SNP IDs, three symbols which are “comma”, “space”, and “line” can be used to separate SNPs. The maximal input SNP IDs is limited to fifty.

2.2 Example: A complete operation using SNP ID to analyze the SNP-RFLPing

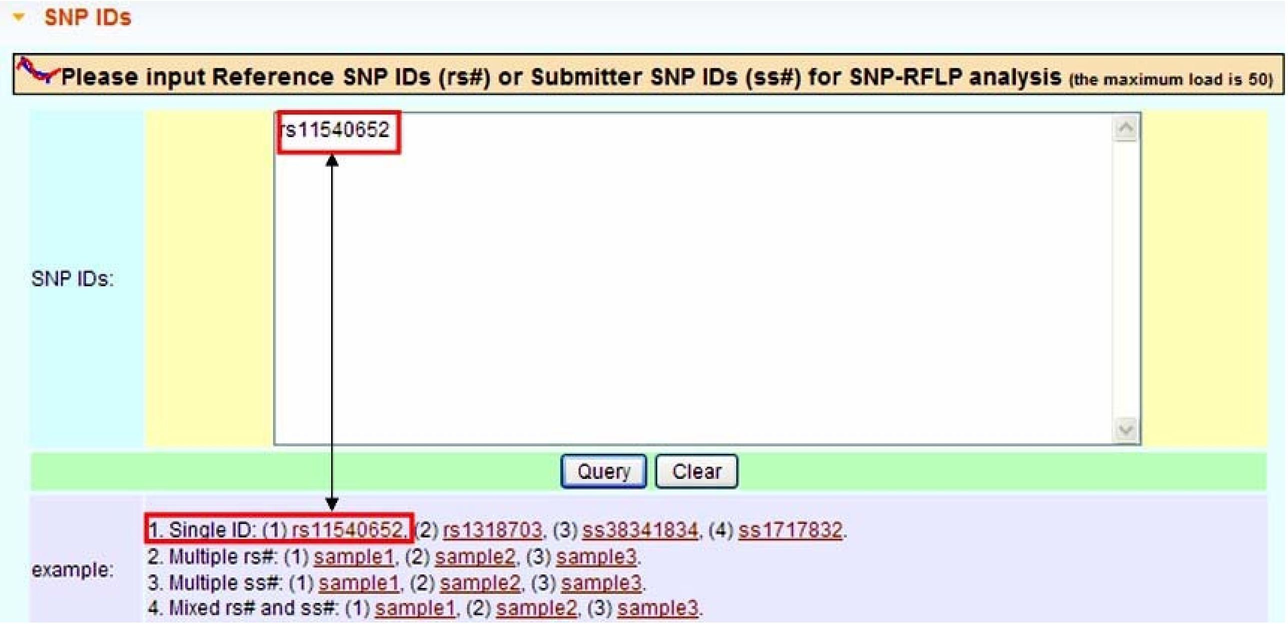
2.2.1 Example (1): SNP IDs input

Fig. 2-1 and Fig. 2-2 show a single ID input for rs# and ss#, respectively.

Fig. 2-3, Fig. 2-4 and Fig. 2-5 show the multiple rs# input with the “comma” symbol, “space”, and “line” (press the “Enter key” in keyboard) to separate, respectively.

Fig. 2-6, Fig. 2-7 and Fig. 2-8 show the multiple ss# input with comma symbol, space, and line to separate, respectively.

Fig. 2-9, Fig. 2-10 and Fig. 2-11 show the multiple rs# and ss# input with comma symbol, space, and line to separate, respectively.



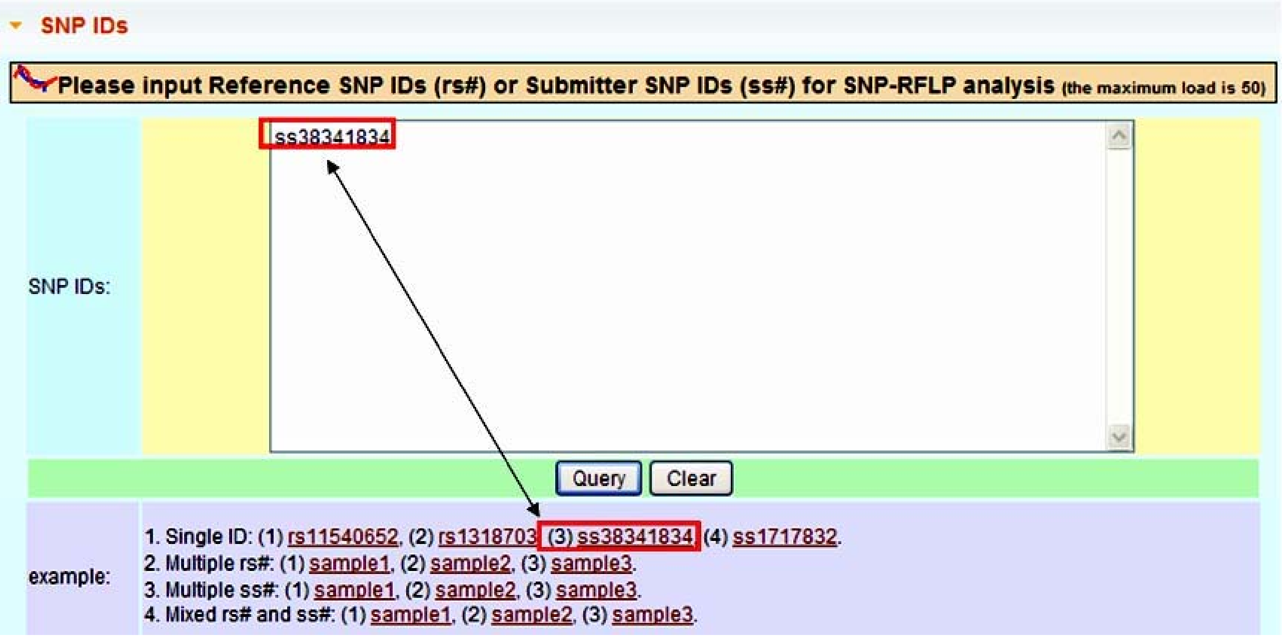
SNP IDs

Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs: rs11540652

example:
1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-1. Single ID input with rs#.

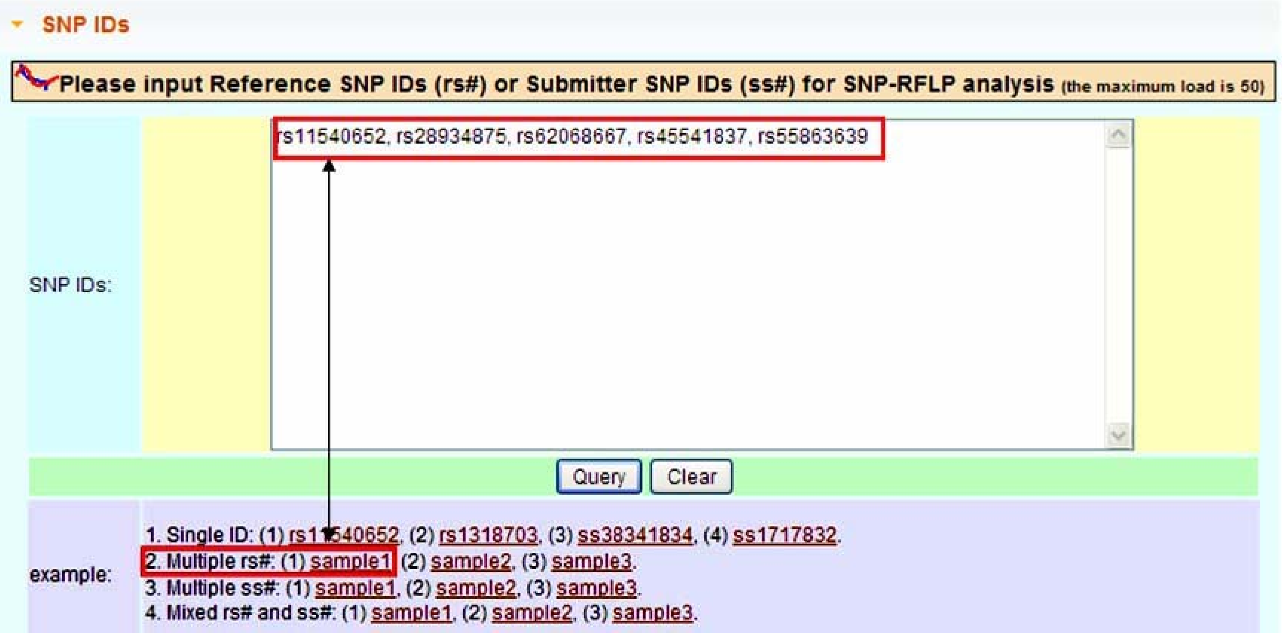


Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs: ss38341834

example:
1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-2. Single ID input with ss#.

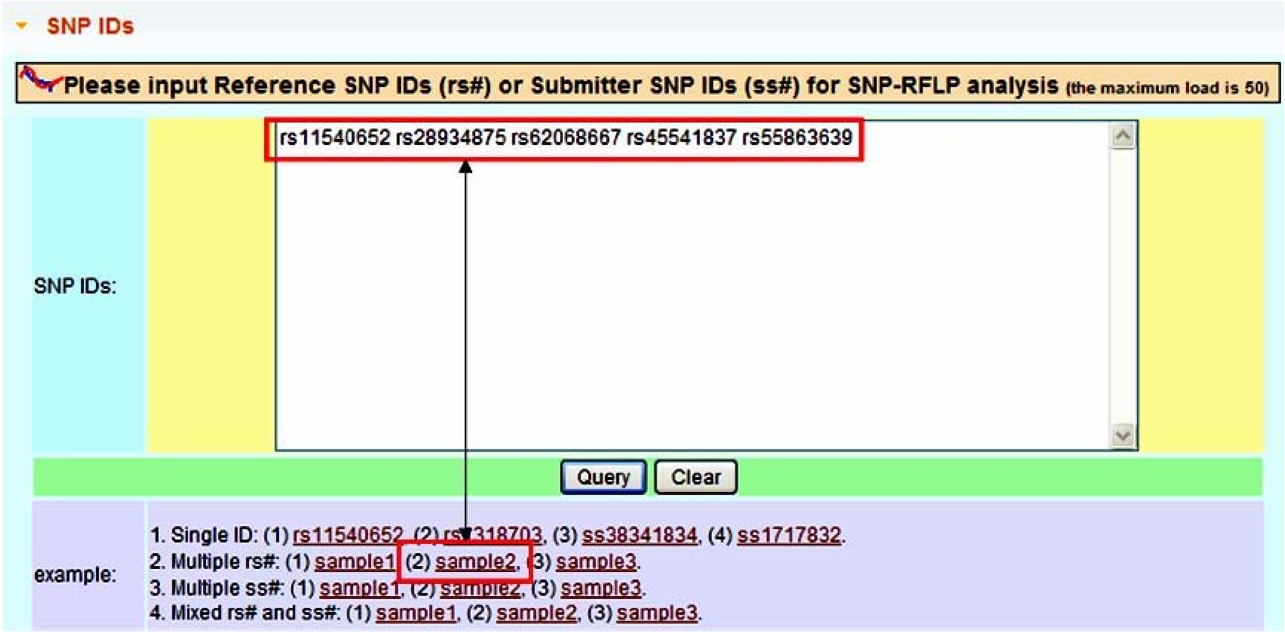


Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs: rs11540652, rs28934875, rs62068667, rs45541837, rs55863639

example:
1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-3. Multiple rs# input with the comma symbols to separate.

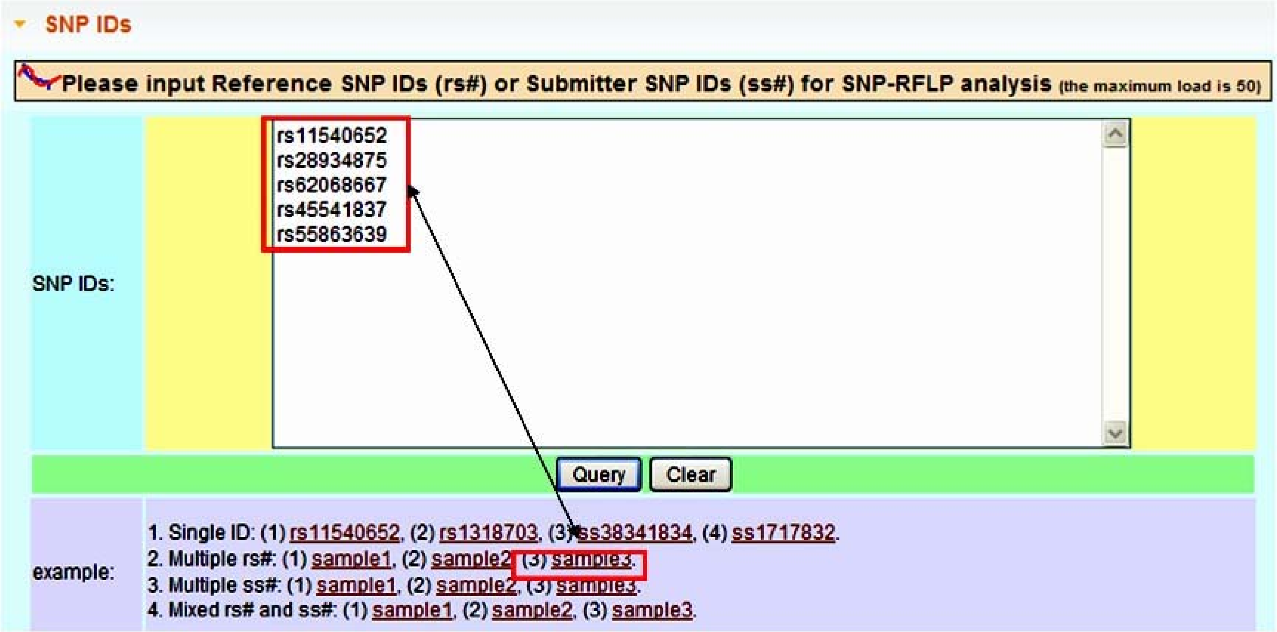


Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs: rs11540652 rs28934875 rs62068667 rs45541837 rs55863639

example:
1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-4. Multiple rs# input with space to separate.

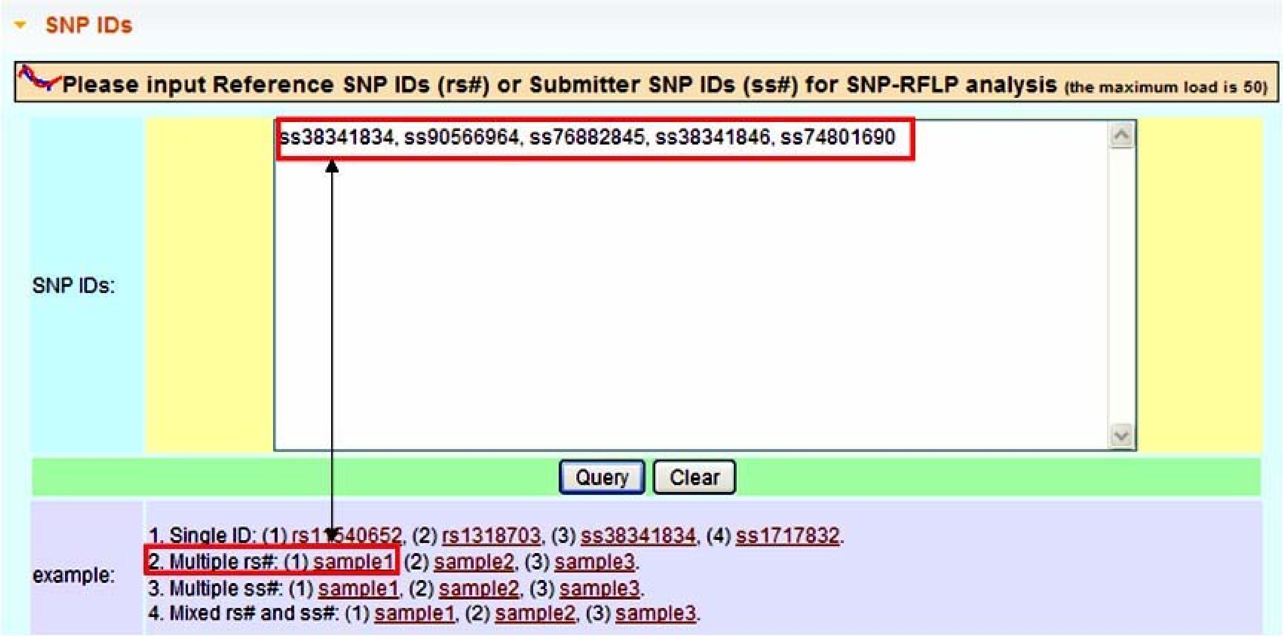


Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs:
rs11540652
rs28934875
rs62068667
rs45541837
rs55863639

example:
1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-5. Multiple rs# input with lines (press the enter key of keyboard) to separate.



SNP IDs

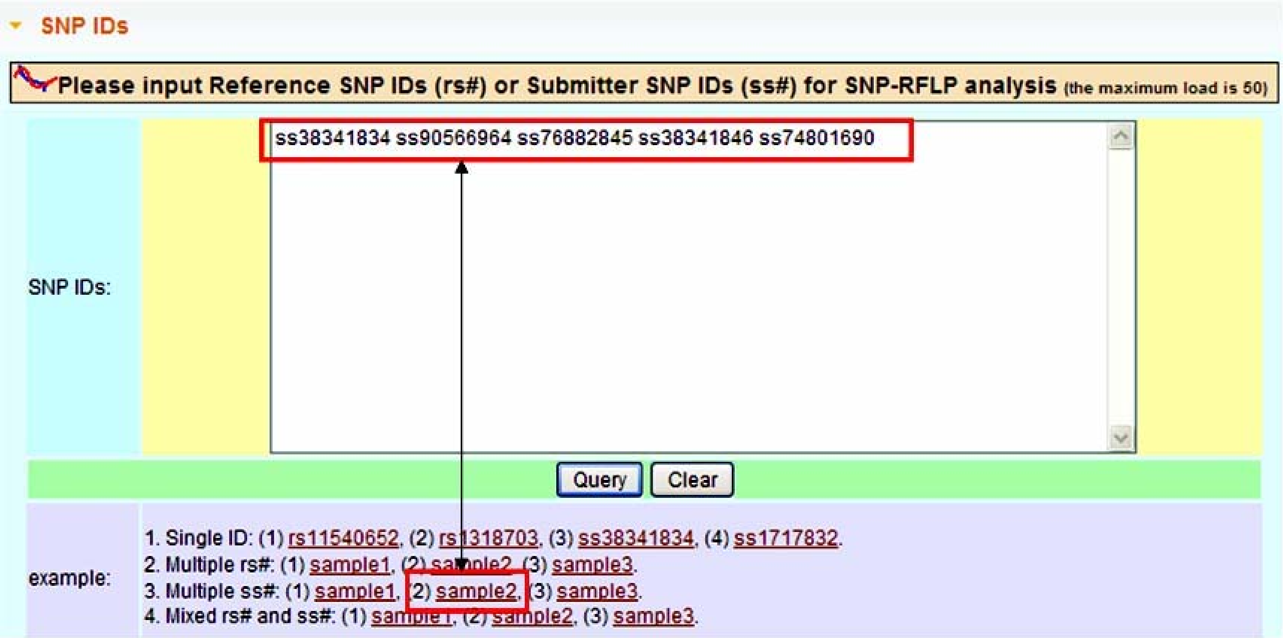
Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs: ss38341834, ss90566964, ss76882845, ss38341846, ss74801690

example:

1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-6. Multiple ss# input with the comma symbols to separate.



SNP IDs

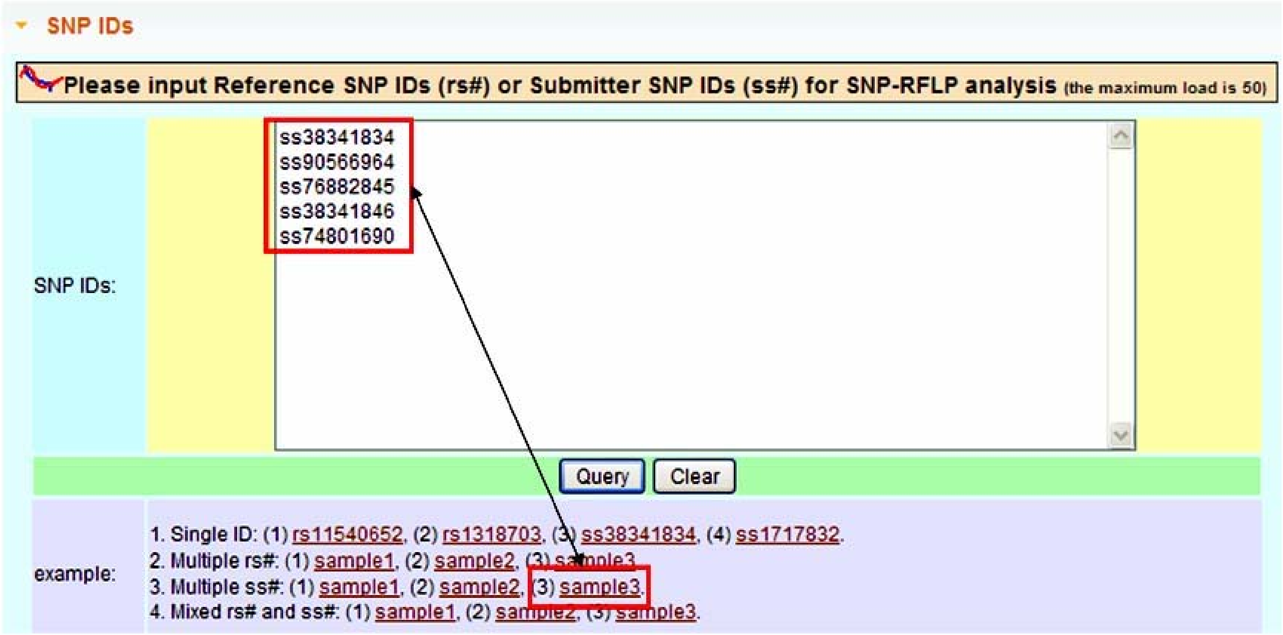
Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs: ss38341834 ss90566964 ss76882845 ss38341846 ss74801690

example:

1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-7. Multiple ss# input with space to separate.

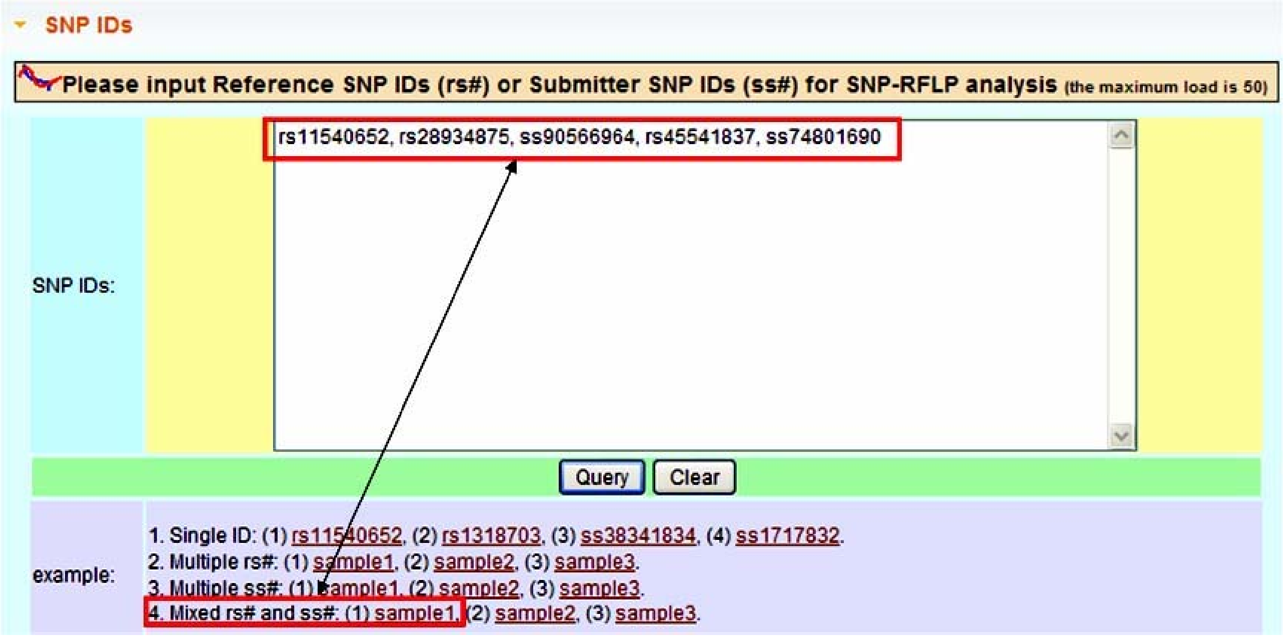


Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs:
```
ss38341834
ss90566964
ss76882845
ss38341846
ss74801690
```

example:
1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-8. Multiple ss# input with lines (press the enter key of keyboard) to separate.

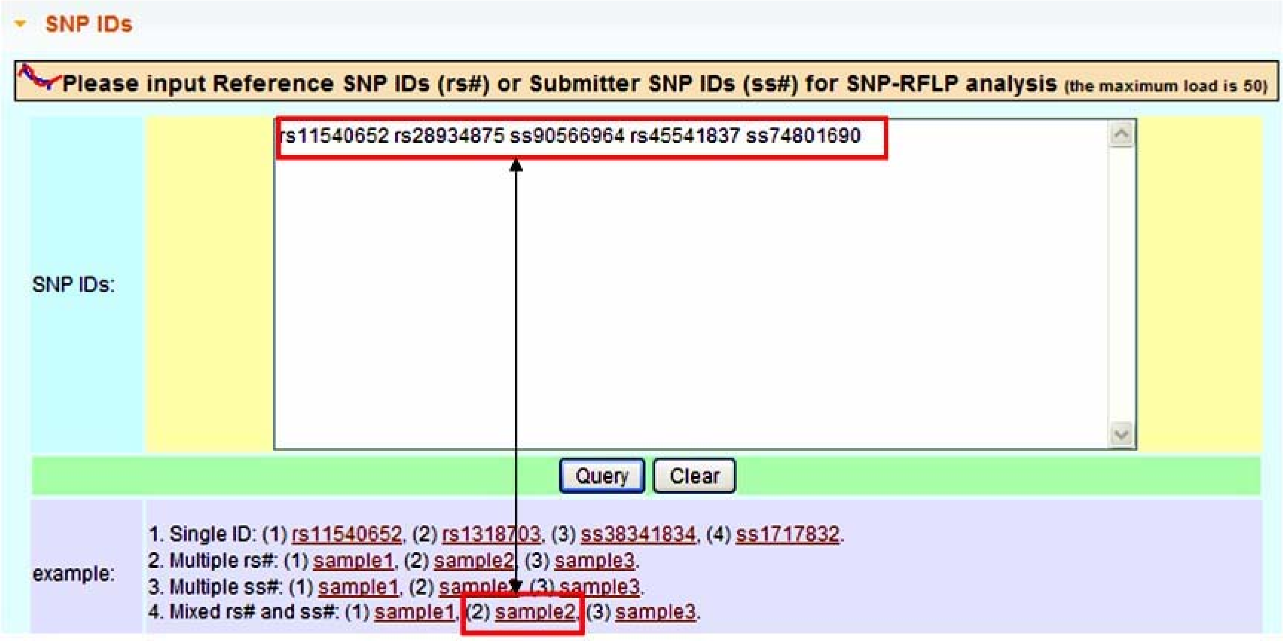


Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs: rs11540652, rs28934875, ss90566964, rs45541837, ss74801690

example:
1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-9. Multiple rs# and ss# input with the comma symbols to separate.

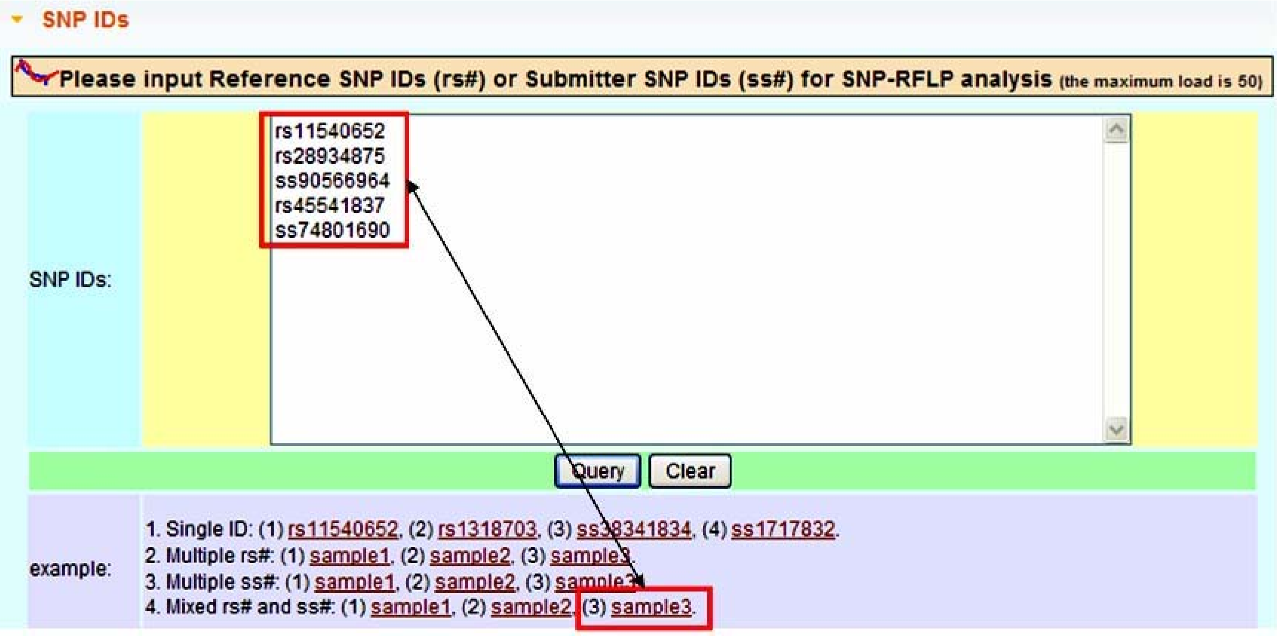


Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

SNP IDs: rs11540652 rs28934875 ss90566964 rs45541837 ss74801690

example:
1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-10. Multiple rs# and ss# input with space to separate.



Please input Reference SNP IDs (rs#) or Submitter SNP IDs (ss#) for SNP-RFLP analysis (the maximum load is 50)

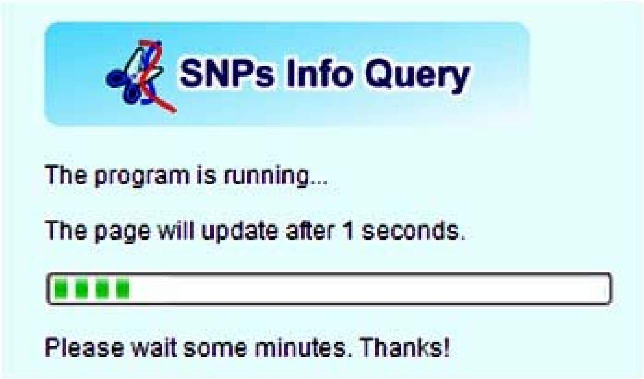
SNP IDs:
rs11540652
rs28934875
ss90566964
rs45541837
ss74801690

example:
1. Single ID: (1) rs11540652, (2) rs1318703, (3) ss38341834, (4) ss1717832.
2. Multiple rs#: (1) sample1, (2) sample2, (3) sample3.
3. Multiple ss#: (1) sample1, (2) sample2, (3) sample3.
4. Mixed rs# and ss#: (1) sample1, (2) sample2, (3) sample3.

Fig. 2-11. Multiple rs# and ss# mixed input with lines (press the enter key of keyboard) to separate.

2.2.2 Example (2): Query SNPs information

After inputting SNP ID, users can click the “Query” button to query SNPs information. Users will see the waiting page (Fig. 2-12) for a moment and the result for SNP information will appear (Fig. 2-13). The result for SNP information includes: reference SNP ID (rs#), organism, short SNP flanking sequence (52 bp in total), chromosome information, clinical/LSDB associated, PubMed citing, SNP gene, sequence viewer, SNP 3D, OMIM, heterozygosity frequency, validated and genotype data, such as NCBI Entrez SNP. Furthermore, HGVS Names are provided for *Homo sapiens*.

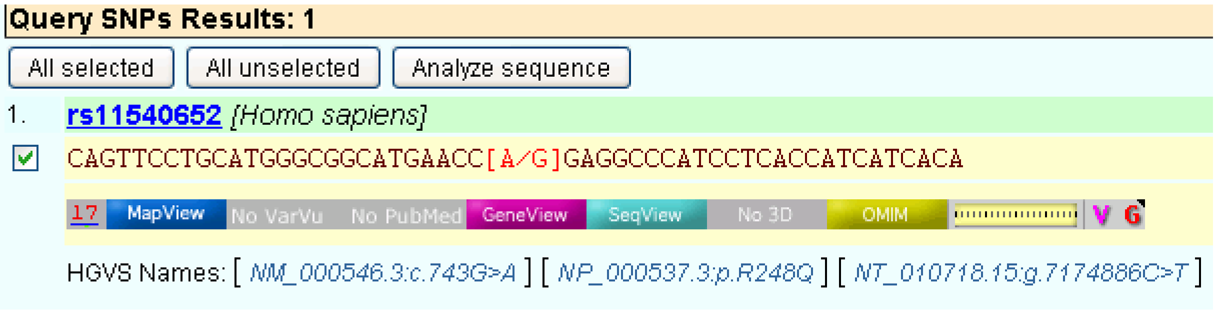

SNPs Info Query

The program is running...

The page will update after 1 seconds.

Please wait some minutes. Thanks!

Fig. 2-12. Waiting for the SNP information query.


Query SNPs Results: 1
All selected
All unselected
Analyze sequence
17
MapView
No VarVu
No PubMed
GeneView
SeqView
No 3D
OMIM
V
G

1. rs11540652 [Homo sapiens]

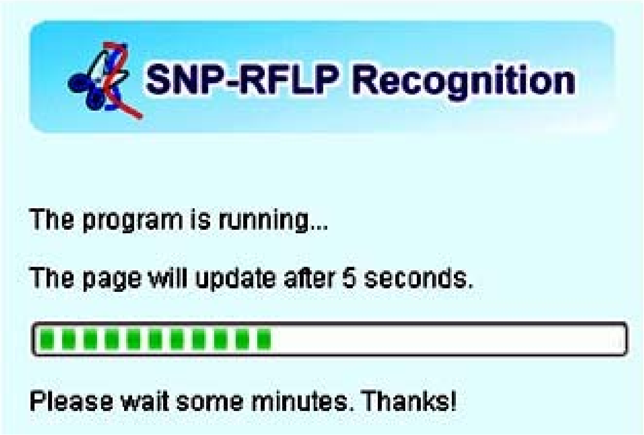
☑ CAGTTCCTGCATGGGCGGCATGAACC[A/G]GAGGCCCATCCTCACCATCATCACA

HGVS Names: [NM_000546.3:c.743G>A] [NP_000537.3:p.R248Q] [NT_010718.15:g.7174886C>T]

Fig. 2-13. The result for SNP information.

2.2.3 Example (3): SNP-RFLP analysis

Subsequently, user can check SNP for further analysis. In this example, a SNP rs11540652 is chosen to check (Fig. 2-13) and users can click the "Analyze sequence" button to execute SNP-RFLP function. When the "Analyze sequence" button is clicked, a waiting page will appear (Fig. 2-14).



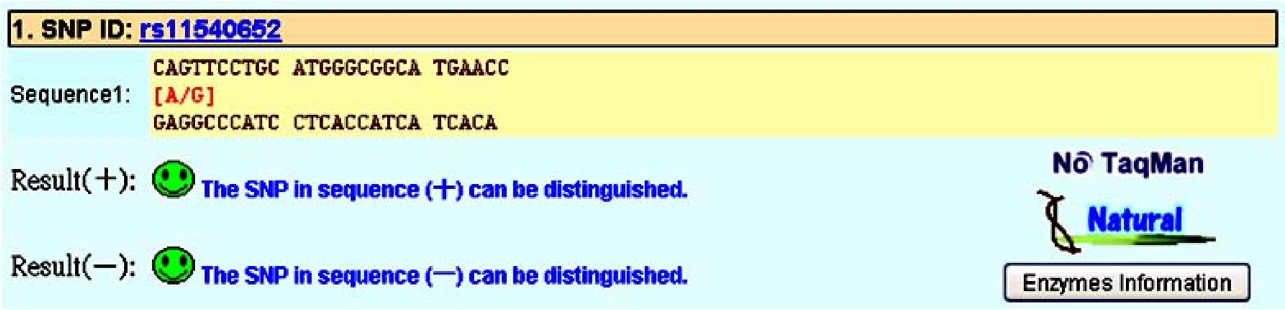
The program is running...

The page will update after 5 seconds.

Please wait some minutes. Thanks!

Fig. 2-14. Waiting for the SNP-RFLP recognition.

After the waiting page is disappeared, the result for SNP-RFLP recognition is appearing (Fig. 2-15). The result information of SNP-RFLP includes SNP ID (rs#), short SNP flanking sequence (52 bp in total), result for sense strand and anti-sense strand whether recognized by restriction enzymes or not, TaqMan, primer set, and enzymes information. The SNP ID is hyperlinked to NCBI Entrez SNP for further information. In this SNP (rs11540652), its TaqMan probe is not provided in both NCBI and SNP500Cancer and the natural primer set (*i.e.*, *the primers are the regular primers and the RFLP restriction enzymes are available*) is available.



1. SNP ID: rs11540652

Sequence1: CAGTTCCTGC ATGGGCGGCA TGAACC
[A/G]
GAGGCCCATC CTCACCATCA TCACA

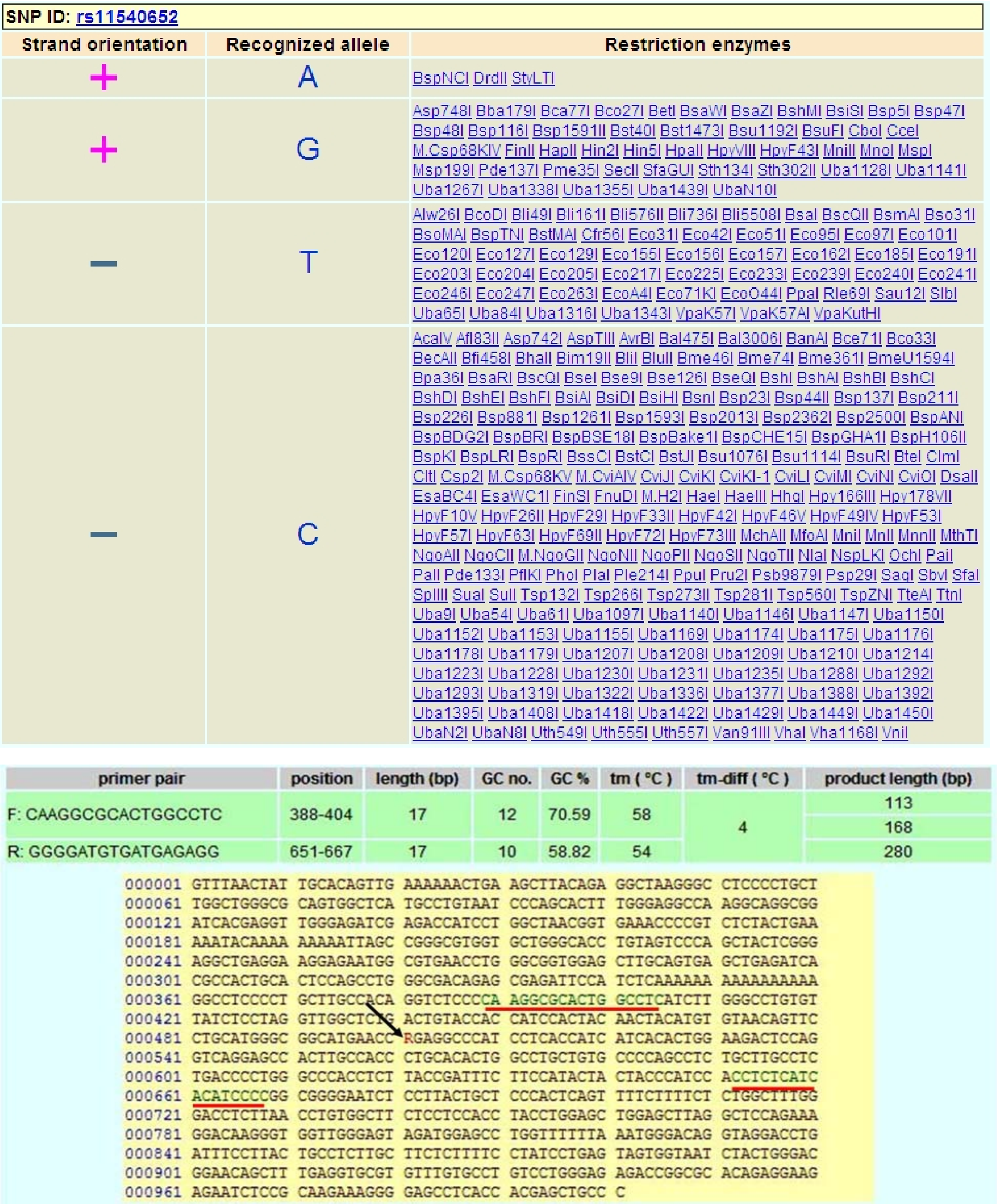
Result(+): The SNP in sequence (+) can be distinguished.

Result(−): The SNP in sequence (−) can be distinguished.

Fig. 2-15. The result for SNP-RFLP recognition. + is the sense strand and – is the antisense strand. Smile symbol means that the RFLP enzymes are available in the SNP-RFLPing 2 analysis.

2.2.4 Example (4): Natural primers information acquiring

Clicking the icon of "Natural", the natural primer set information is appearing (Fig. 2-16). The natural primer set information contains SNP ID, strand orientation, recognized allele, restriction enzymes, natural primer pair information, and natural primers visualization.



SNP ID: rs11540652

Strand orientation	Recognized allele	Restriction enzymes
+	A	BspNCI DrdII StyLTI
+	G	Asp748I Bba179I Bca77I Bco27I BetI BsaWI BsaZI BshMI BsiSI Bsp5I Bsp47I Bsp48I Bsp116I Bsp1591II Bst40I Bst1473I Bsu1192I BsuFI CboI CceI M.Csp68KIV FinII HapII Hin2I Hin5I HpaII HpyVIII HpyF43I MniII MnoI MspI Msp199I Pde137I Pme35I SecII SfaGUI Sth134I Sth302II Uba1128I Uba1141I Uba1267I Uba1338I Uba1355I Uba1439I UbaN10I
–	T	Alw26I BcoDI Bli49I Bli161I Bli576II Bli736I Bli5508I BsaI BscQII BsmAI Bso31I BsoMAI BspTNI BstMAI Cfr56I Eco31I Eco42I Eco51I Eco95I Eco97I Eco101I Eco120I Eco127I Eco129I Eco155I Eco156I Eco157I Eco162I Eco185I Eco191I Eco203I Eco204I Eco205I Eco217I Eco225I Eco233I Eco239I Eco240I Eco241I Eco246I Eco247I Eco263I EcoA4I Eco71KI EcoO44I PpaI Rle69I Sau12I SlbI Uba65I Uba84I Uba1316I Uba1343I VpaK57I VpaK57AI VpaKutHI
–	C	AcaIV Afl83II Asp742I AspTIII AvrBI Bal475I Bal3006I BanAI Bce71I Bco33I BecAII Bfi458I BhaII Bim19II BliI BluII Bme46I Bme74I Bme361I BmeU1594I Bpa36I BsaRI BscQI BseI Bse9I Bse126I BseQI BshI BshAI BshBI BshCI BshDI BshEI BshFI BsiAI BsiDI BsiHI BsnI Bsp23I Bsp44II Bsp137I Bsp211I Bsp226I Bsp881I Bsp1261I Bsp1593I Bsp2013I Bsp2362I Bsp2500I BspANI BspBDG2I BspBRI BspBSE18I BspBake1I BspCHE15I BspGHA1I BspH106II BspKI BspLRI BspRI BssCI BstCI BstJI Bsu1076I Bsu1114I BsuRI BteI ClmI CltI Csp2I M.Csp68KV M.CviAIV CviJI CviKI CviKI-1 CviLI CviMI CviNI CviOI DsaII EsaBC4I EsaWC1I FinSI FnuDI M.H2I HaeI HaeIII HhgI Hpy166III Hpy178VII HpyF10V HpyF26II HpyF29I HpyF33II HpyF42I HpyF46V HpyF49IV HpyF53I HpyF57I HpyF63I HpyF69II HpyF72I HpyF73III MchAII MfoAI MniI MnlI MnnII MthTI NgoAII NgoCII M.NgoGII NgoNII NgoPII NgoSII NgoTII NlaI NspLKI OchI PaiI PalI Pde133I PflKI PhoI PlaI Ple214I PpuI Pru2I Psb9879I Psp29I SaqI SbvI SfaI SplIII SuaI SulI Tsp132I Tsp266I Tsp273II Tsp281I Tsp560I TspZNI TteAI TtnI Uba9I Uba54I Uba61I Uba1097I Uba1140I Uba1146I Uba1147I Uba1150I Uba1152I Uba1153I Uba1155I Uba1169I Uba1174I Uba1175I Uba1176I Uba1178I Uba1179I Uba1207I Uba1208I Uba1209I Uba1210I Uba1214I Uba1223I Uba1228I Uba1230I Uba1231I Uba1235I Uba1288I Uba1292I Uba1293I Uba1319I Uba1322I Uba1336I Uba1377I Uba1388I Uba1392I Uba1395I Uba1408I Uba1418I Uba1422I Uba1429I Uba1449I Uba1450I UbaN2I UbaN8I Uth549I Uth555I Uth557I Van91II VhaI Vha1168I VniI

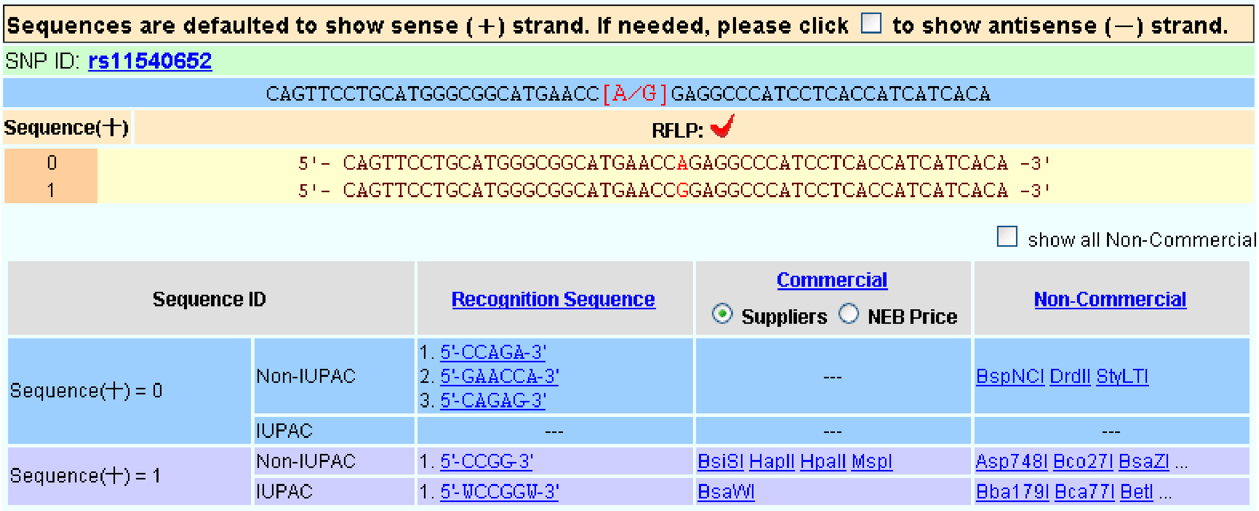
primer pair	position	length (bp)	GC no.	GC %	tm (°C)	tm-diff (°C)	product length (bp)
F: CAAGGCGCACTGGCCTC	388-404	17	12	70.59	58	4	113 168
R: GGGGATGTGATGAGAGG	651-667	17	10	58.82	54		280

```
000001 GTTTAACTAT TGCACAGTTG AAAAAACTGA AGCTTACAGA GGCTAAGGGC CTCCCCTGCT
000061 TGGCTGGGCG CAGTGGCTCA TGCCTGTAAT CCCAGCACTT TGGGAGGCCA AGGCAGGCGG
000121 ATCACGAGGT TGGGAGATCG AGACCATCCT GGCTAACGGT GAAACCCCGT CTCTACTGAA
000181 AAATACAAAA AAAAATTAGC CGGGCGTGGT GCTGGGCACC TGTAGTCCCA GCTACTCGGG
000241 AGGCTGAGGA AGGAGAATGG CGTGAACCTG GGCGGTGGAG CTTGCAGTGA GCTGAGATCA
000301 CGCCACTGCA CTCCAGCCTG GGCGACAGAG CGAGATTCCA TCTCAAAAAA AAAAAAAAAA
000361 GGCCTCCCCT GCTTGCCACA GGTCTCCCCA AGGCGCACTG GCCTCATCTT GGGCCTGTGT
000421 TATCTCCTAG GTTGGCTCTG ACTGTACCAC CATCCACTAC AACTACATGT GTAACAGTTC
000481 CTGCATGGGC GGCATGAACC RGAGGCCCAT CCTCACCATC ATCACACTGG AAGACTCCAG
000541 GTCAGGAGCC ACTTGCCACC CTGCACACTG GCCTGCTGTG CCCCAGCCTC TGCTTGCCTC
000601 TGACCCCTGG GCCCACCTCT TACCGATTTC TTCCATACTA CTACCCATCC ACCTCTCATC
000661 ACATCCCCGG CGGGGAATCT CCTTACTGCT CCCACTCAGT TTTCTTTTCT CTGGCTTTGG
000721 GACCTCTTAA CCTGTGGCTT CTCCTCCACC TACCTGGAGC TGGAGCTTAG GCTCCAGAAA
000781 GGACAAGGGT GGTTGGGAGT AGATGGAGCC TGGTTTTTTA AATGGGACAG GTAGGACCTG
000841 ATTTCCTTAC TGCCTCTTGC TTCTCTTTTC CTATCCTGAG TAGTGGTAAT CTACTGGGAC
000901 GGAACAGCTT TGAGGTGCGT GTTTGTGCCT GTCCTGGGAG AGACCGGCGC ACAGAGGAAG
000961 AGAATCTCCG CAAGAAAGGG GAGCCTCACC ACGAGCTGCC C
```

Fig. 2-16. The restriction enzymes and natural primer set information for rs11540652. Both forward and reverse primers are underlined with red color and the SNP is indicated by an arrow line. + is the sense strand and – is the antisense strand.

2.2.5 Example (5): Available restriction enzymes request

In order to further get full enzymes information, users can click the “Enzymes Information” button. Fig. 2-17 shows only the available “sense” strand restriction enzymes information. When the checkbox for anti-sense strand is checked, the restriction enzymes information for both the available “sense” and “anti-sense” strands are shown in Fig. 2-18.



Sequences are defaulted to show sense (+) strand. If needed, please click ☐ to show antisense (—) strand.

SNP ID: rs11540652

CAGTTCCTGCATGGGCGGCATGAACC[A/G]GAGGCCCATCCTCACCATCATCACA

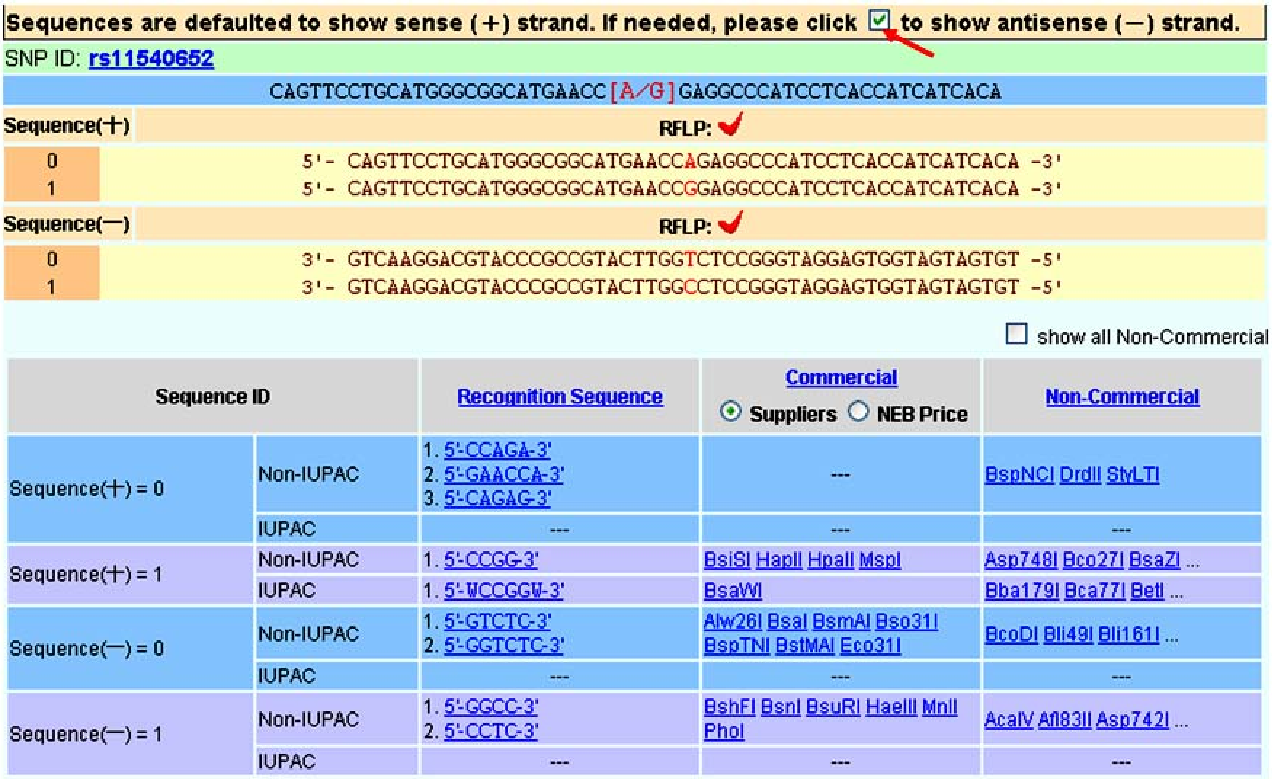
Sequence(+) **RFLP:** ✓

0 5'- CAGTTCCTGCATGGGCGGCATGAACCAGAGGCCCATCCTCACCATCATCACA -3'
1 5'- CAGTTCCTGCATGGGCGGCATGAACCGGAGGCCCATCCTCACCATCATCACA -3'

☐ show all Non-Commercial

Sequence ID		**Recognition Sequence**	**Commercial** ● **Suppliers** ○ **NEB Price**	**Non-Commercial**
Sequence(+) = 0	Non-IUPAC	1. 5'-CCAGA-3' 2. 5'-GAACCA-3' 3. 5'-CAGAG-3'	---	BspNCI DrdII StyLTI
	IUPAC	---	---	---
Sequence(+) = 1	Non-IUPAC	1. 5'-CCGG-3'	BsiSI HapII HpaII MspI	Asp748I Bco27I BsaZI ...
	IUPAC	1. 5'-WCCGGW-3'	BsaWI	Bba179I Bca77I BetI ...

Fig. 2-17. The available sense strand restriction enzymes information for rs11540652. + is the sense strand and – is the antisense strand. 0 and 1 indicate the alternative nucleotide for the same SNP.



Sequences are defaulted to show sense (+) strand. If needed, please click ☑ to show antisense (—) strand.

SNP ID: rs11540652

CAGTTCCTGCATGGGCGGCATGAACC[A/G]GAGGCCCATCCTCACCATCATCACA

Sequence(+) **RFLP:** ✓

0 5'- CAGTTCCTGCATGGGCGGCATGAACCAGAGGCCCATCCTCACCATCATCACA -3'
1 5'- CAGTTCCTGCATGGGCGGCATGAACCGGAGGCCCATCCTCACCATCATCACA -3'

Sequence(—) **RFLP:** ✓

0 3'- GTCAAGGACGTACCCGCCGTACTTGGTCTCCGGGTAGGAGTGGTAGTAGTGT -5'
1 3'- GTCAAGGACGTACCCGCCGTACTTGGCCTCCGGGTAGGAGTGGTAGTAGTGT -5'

☐ show all Non-Commercial

Sequence ID		**Recognition Sequence**	**Commercial** ● **Suppliers** ○ **NEB Price**	**Non-Commercial**
Sequence(+) = 0	Non-IUPAC	1. 5'-CCAGA-3' 2. 5'-GAACCA-3' 3. 5'-CAGAG-3'	---	BspNCI DrdII StyLTI
	IUPAC	---	---	---
Sequence(+) = 1	Non-IUPAC	1. 5'-CCGG-3'	BsiSI HapII HpaII MspI	Asp748I Bco27I BsaZI ...
	IUPAC	1. 5'-WCCGGW-3'	BsaWI	Bba179I Bca77I BetI ...
Sequence(—) = 0	Non-IUPAC	1. 5'-GTCTC-3' 2. 5'-GGTCTC-3'	Alw26I BsaI BsmAI Bso31I BspTNI BstMAI Eco31I	BcoDI Bli49I Bli161I ...
	IUPAC	---	---	---
Sequence(—) = 1	Non-IUPAC	1. 5'-GGCC-3' 2. 5'-CCTC-3'	BshFI BsnI BsuRI HaeIII MnlI PhoI	AcaIV Afl83II Asp742I ...
	IUPAC	---	---	---

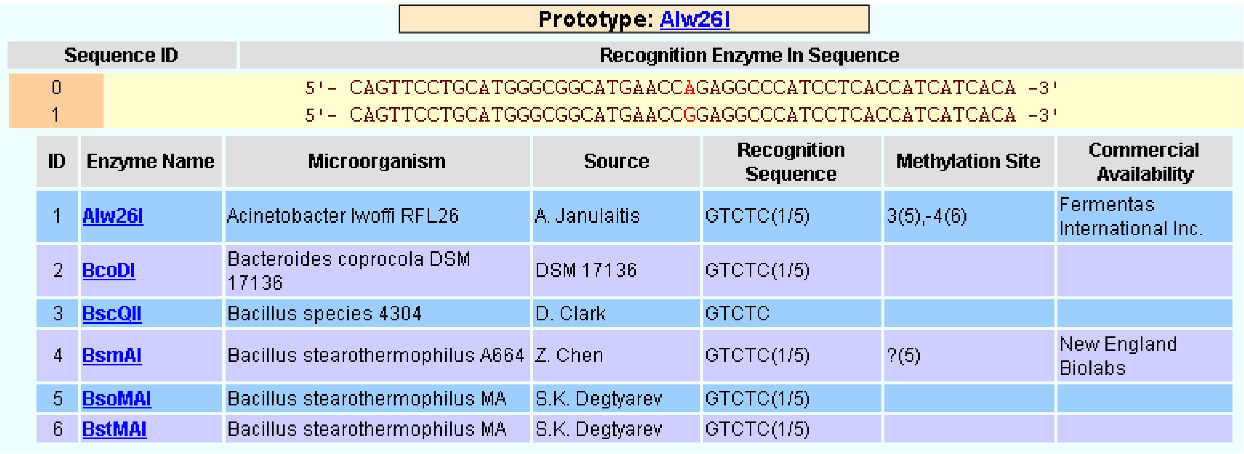
Fig. 2-18. The available sense and anti-sense strand restriction enzyme information for rs11540652 when the checkbox for anti-sense strand (indicated by a red arrow line) is checked. + is the sense strand and – is the antisense strand. 0 and 1 indicate the alternative nucleotide for the same SNP.

2.2.6 Example (6): Further restriction enzyme information description

The enzymes information contains: (1) “Sequence ID” for sense and anti-sense strands, (2) “Recognition Sequence”, and (3) “Commercial” and “Non-Commercial” enzymes.

For “Recognition Sequence”, this system provides the hyperlink for the same recognition sequence (Fig. 2-19). When the hyperlink for “Enzyme Name” is clicked, the search results for this enzyme name in NEW ENGLANE BioLabs (NEB) will be output (Fig. 2-20).

For “Commercial” enzymes, two options are available: one is ‘Suppliers’ and the other is ‘NEB Price’. When users select the ‘Suppliers’, the following enzyme hyperlinks will link to REBASE Suppliers (Fig. 2-21). When users select the ‘NEB Price’, the following enzymes hyperlink will link to price manager for Restriction Endonucleases which is our locally built-in database, all enzyme prices come from the price for NEB (Fig. 2-22 and Fig. 2-23). For “Non-Commercial” enzymes, enzymes hyperlink will be link to NCBI PubMed to search related documents (Fig. 2-24).

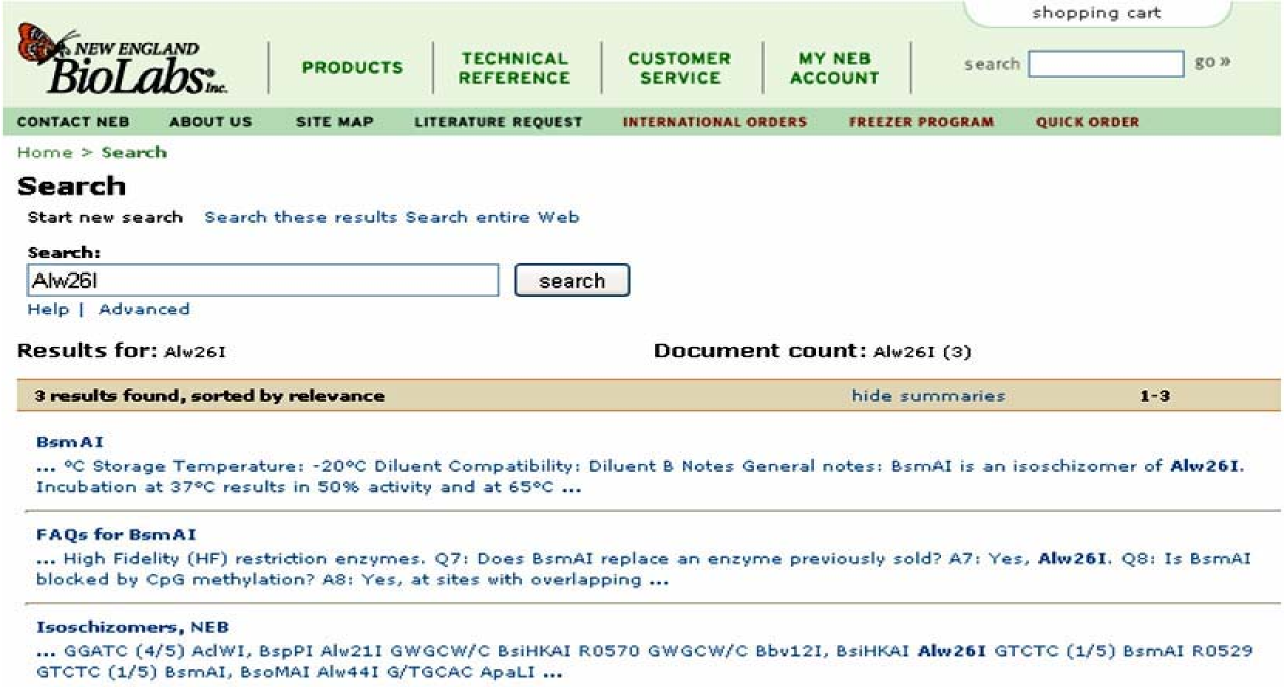


Prototype: Alw26I

Sequence ID	Recognition Enzyme In Sequence
0	5'- CAGTTCCTGCATGGGCGGCATGAACCAGAGGCCCATCCTCACCATCATCACA -3'
1	5'- CAGTTCCTGCATGGGCGGCATGAACCGGAGGCCCATCCTCACCATCATCACA -3'

ID	Enzyme Name	Microorganism	Source	Recognition Sequence	Methylation Site	Commercial Availability
1	Alw26I	Acinetobacter lwoffi RFL26	A. Janulaitis	GTCTC(1/5)	3(5),-4(6)	Fermentas International Inc.
2	BcoDI	Bacteroides coprocola DSM 17136	DSM 17136	GTCTC(1/5)		
3	BscQII	Bacillus species 4304	D. Clark	GTCTC		
4	BsmAI	Bacillus stearothermophilus A664	Z. Chen	GTCTC(1/5)	?(5)	New England Biolabs
5	BsoMAI	Bacillus stearothermophilus MA	S.K. Degtyarev	GTCTC(1/5)		
6	BstMAI	Bacillus stearothermophilus MA	S.K. Degtyarev	GTCTC(1/5)		

Fig. 2-19. The different enzyme names with the same recognition sequence ‘GTCTC’.



Search

Start new search   Search these results Search entire Web

Search: Alw26I   search

Help | Advanced

Results for: Alw26I   **Document count:** Alw26I (3)

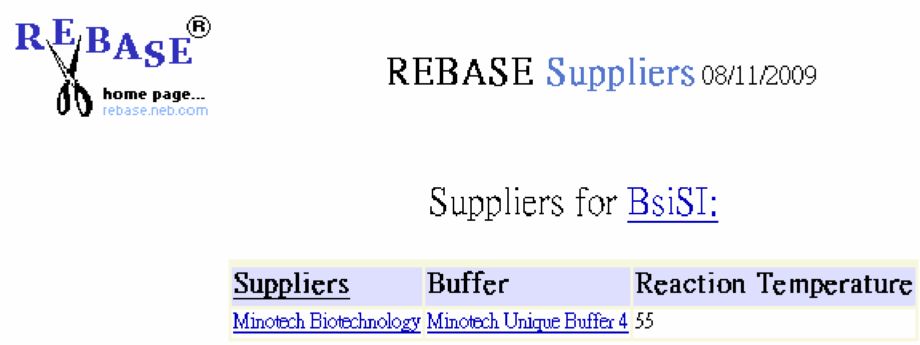
3 results found, sorted by relevance   hide summaries   **1-3**

BsmAI
... °C Storage Temperature: -20°C Diluent Compatibility: Diluent B Notes General notes: BsmAI is an isoschizomer of **Alw26I**. Incubation at 37°C results in 50% activity and at 65°C ...

FAQs for BsmAI
... High Fidelity (HF) restriction enzymes. Q7: Does BsmAI replace an enzyme previously sold? A7: Yes, **Alw26I**. Q8: Is BsmAI blocked by CpG methylation? A8: Yes, at sites with overlapping ...

Isoschizomers, NEB
... GGATC (4/5) AclWI, BspPI Alw21I GWGCW/C BsiHKAI R0570 GWGCW/C Bbv12I, BsiHKAI **Alw26I** GTCTC (1/5) BsmAI R0529 GTCTC (1/5) BsmAI, BsoMAI Alw44I G/TGCAC ApaLI ...

Fig. 2-20. The search results for Alw26I in NEW ENGLANE BioLabs.

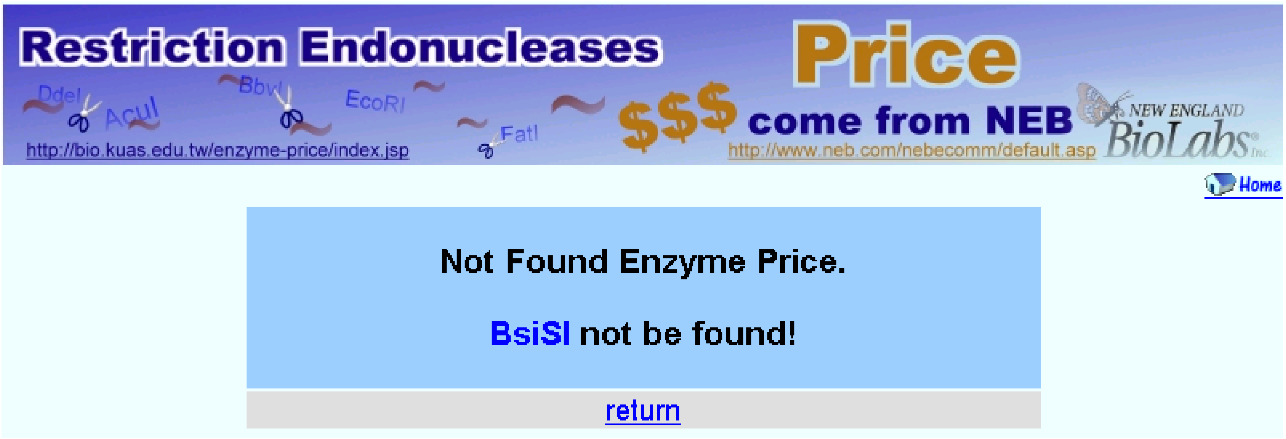

REBASE
home page...
rebase.neb.com
REBASE
Suppliers

REBASE Suppliers 08/11/2009

Suppliers for BsiSI:

Suppliers	Buffer	Reaction Temperature
Minotech Biotechnology	Minotech Unique Buffer 4	55

Fig. 2-21. "Commercial" enzymes option is 'Suppliers'. Clicking the BsiSI enzyme will link to REBASE Suppliers and show the suppliers for BsiSI.

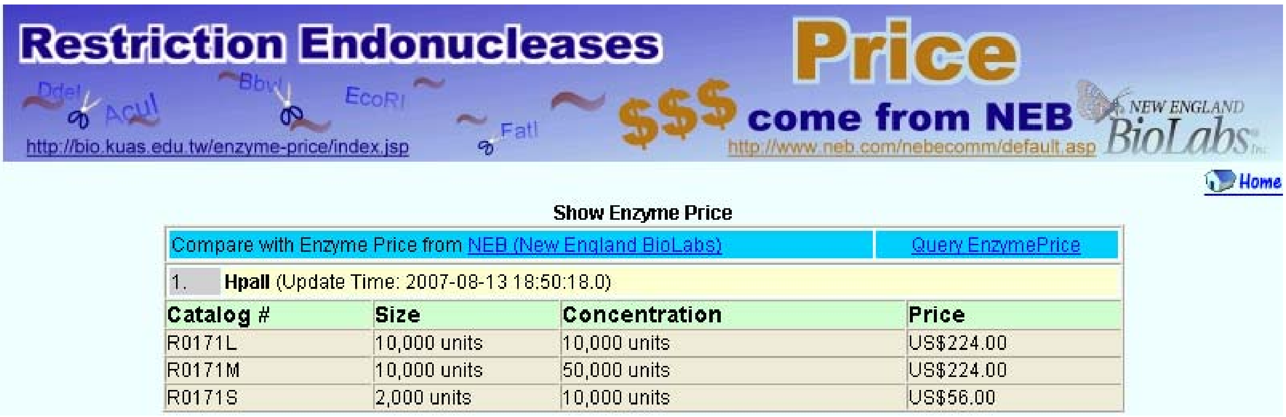

Restriction Endonucleases
Price
$$$ come from NEB
DdeI
AcuI
BbvI
EcoRI
FatI
http://bio.kuas.edu.tw/enzyme-price/index.jsp
http://www.neb.com/nebecomm/default.asp
NEW ENGLAND BioLabs Inc.
Home

Not Found Enzyme Price.

BsiSI not be found!

return

Fig. 2-22. For enzyme BsiSI, no any price is provided.


Restriction Endonucleases
Price
$$$ come from NEB
DdeI
AcuI
BbvI
EcoRI
FatI
http://bio.kuas.edu.tw/enzyme-price/index.jsp
http://www.neb.com/nebecomm/default.asp
NEW ENGLAND BioLabs Inc.
Home

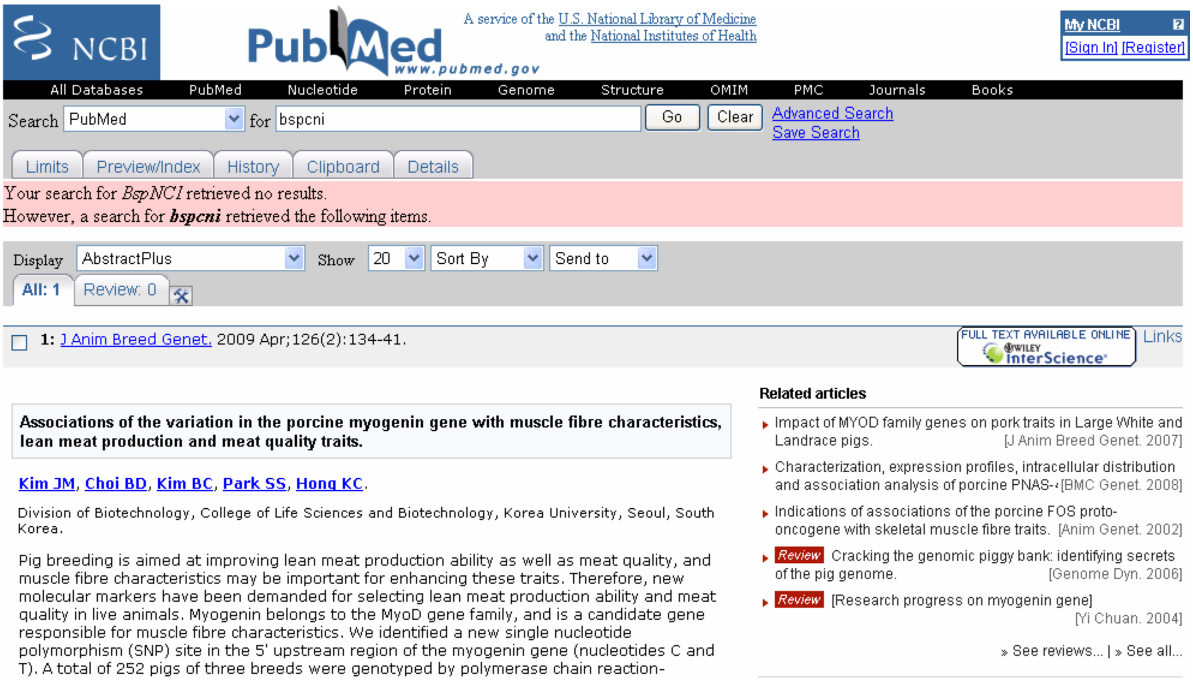
Show Enzyme Price

Compare with Enzyme Price from NEB (New England BioLabs) | Query EnzymePrice

1. **HpaII** (Update Time: 2007-08-13 18:50:18.0)

Catalog #	Size	Concentration	Price
R0171L	10,000 units	10,000 units	US$224.00
R0171M	10,000 units	50,000 units	US$224.00
R0171S	2,000 units	10,000 units	US$56.00

Fig. 2-23. For enzyme HpaII, the price is provided from NEB.


NCBI
PubMed
www.pubmed.gov
A service of the U.S. National Library of Medicine
and the National Institutes of Health
My NCBI
[Sign In] [Register]
All Databases
PubMed
Nucleotide
Protein
Genome
Structure
OMIM
PMC
Journals
Books
Go
Clear
Advanced Search
Save Search
Limits
Preview/Index
History
Clipboard
Details
Sort By
Send to
All: 1
Review: 0
FULL TEXT AVAILABLE ONLINE
WILEY InterScience
Links
Review
» See reviews... | » See all...

Search PubMed for bspcni

Your search for *BspNCI* retrieved no results.
However, a search for ***bspcni*** retrieved the following items.

Display AbstractPlus Show 20

☐ **1:** J Anim Breed Genet. 2009 Apr;126(2):134-41.

Associations of the variation in the porcine myogenin gene with muscle fibre characteristics, lean meat production and meat quality traits.

Kim JM, **Choi BD**, **Kim BC**, **Park SS**, **Hong KC**.

Division of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul, South Korea.

Pig breeding is aimed at improving lean meat production ability as well as meat quality, and muscle fibre characteristics may be important for enhancing these traits. Therefore, new molecular markers have been demanded for selecting lean meat production ability and meat quality in live animals. Myogenin belongs to the MyoD gene family, and is a candidate gene responsible for muscle fibre characteristics. We identified a new single nucleotide polymorphism (SNP) site in the 5' upstream region of the myogenin gene (nucleotides C and T). A total of 252 pigs of three breeds were genotyped by polymerase chain reaction-

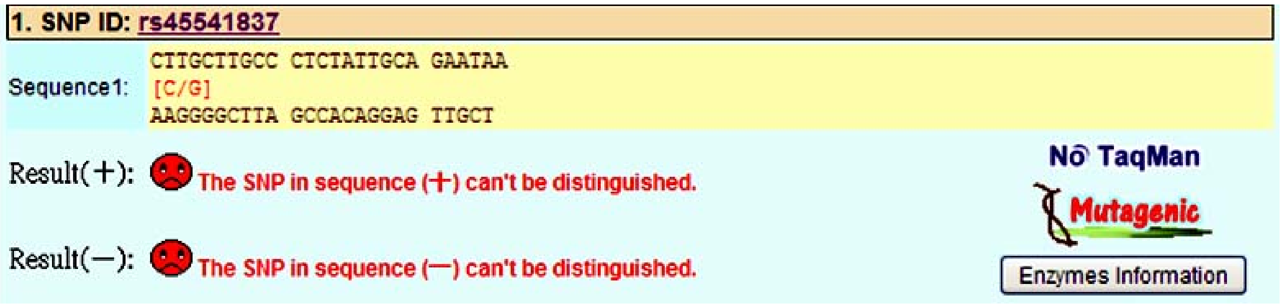
Related articles

- Impact of MYOD family genes on pork traits in Large White and Landrace pigs. [J Anim Breed Genet. 2007]
- Characterization, expression profiles, intracellular distribution and association analysis of porcine PNAS-4 [BMC Genet. 2008]
- Indications of associations of the porcine FOS proto-oncogene with skeletal muscle fibre traits. [Anim Genet. 2002]
- Review Cracking the genomic piggy bank: identifying secrets of the pig genome. [Genome Dyn. 2006]
- Review [Research progress on myogenin gene] [Yi Chuan. 2004]

Fig. 2-24. Clicking the non-Commercial enzymes will be sent to NCBI PubMed to search related documents.

2.2.7 Example (7): Mutagenic primer information acquiring

Some SNPs are not recognized by any restriction enzymes. Alternatively, the mutagenic primer may be designed by the SNP-RFLPing 2 system. When a SNP does not have any restriction enzymes, the icon of “Mutagenic” will appear after SNP-RFLP analysis (Fig. 2-25). Clicking the icon of “Mutagenic”, the mutagenic primer set information is appearing (Fig. 2-26). The mutagenic primer set information contains SNP ID, strand orientation, recognized allele, restriction enzymes, mutagenic primer pair information, and mutagenic primer visualization.



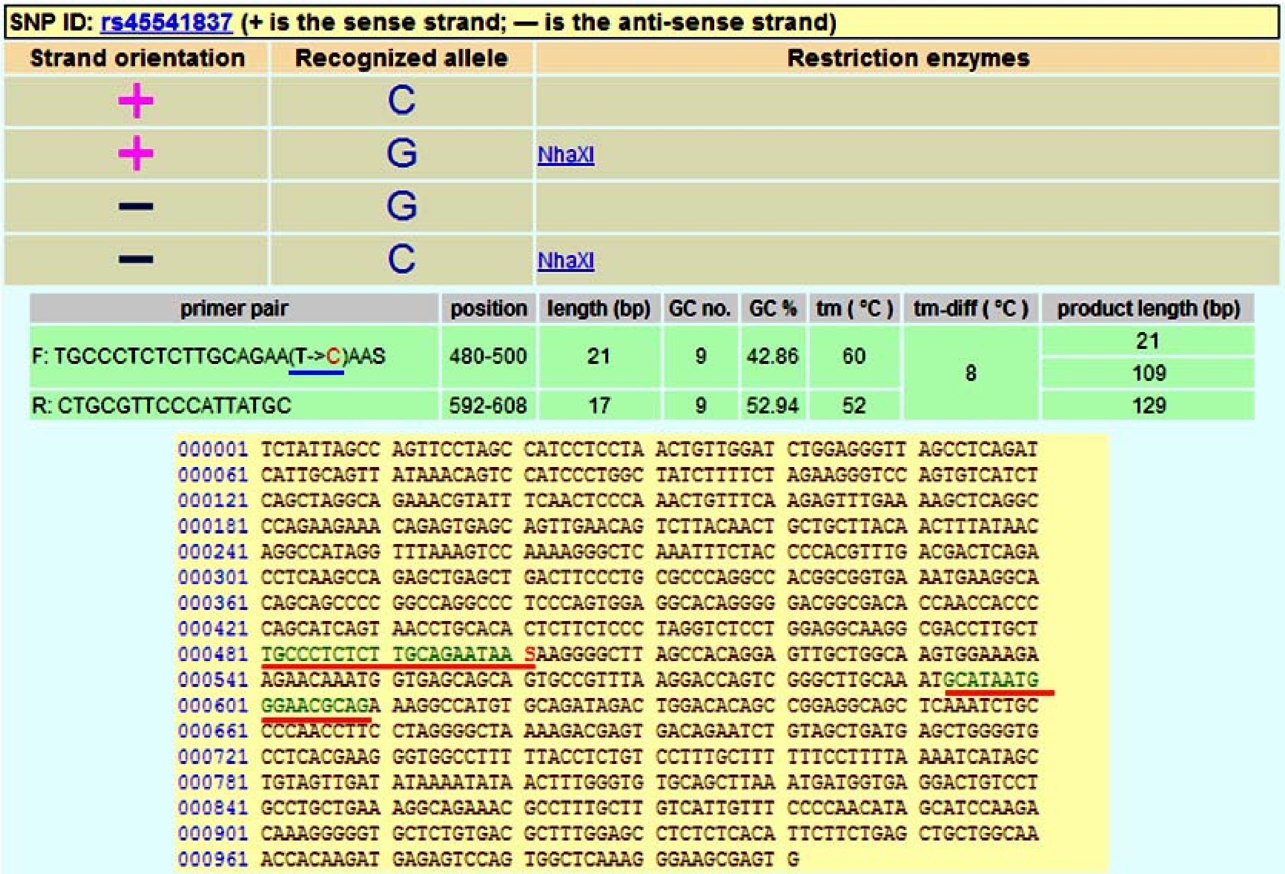
1. SNP ID: rs45541837

Sequence1: CTTGCTTGCC CTCTATTGCA GAATAA [C/G] AAGGGGCTTA GCCACAGGAG TTGCT

Result(+): The SNP in sequence (+) can't be distinguished.

Result(−): The SNP in sequence (−) can't be distinguished.

Fig. 2-25. No any available restriction enzymes for SNP rs45541837, the icon of “Mutagenic” appears in the result of SNP-RFLP analysis.



SNP ID: rs45541837 (+ is the sense strand; — is the anti-sense strand)

Strand orientation	Recognized allele	Restriction enzymes
+	C	
+	G	NhaXI
−	G	
−	C	NhaXI

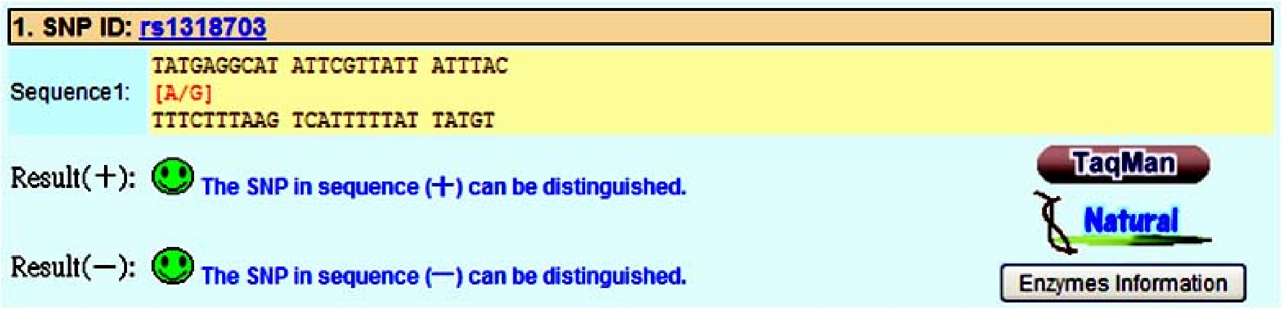
primer pair	position	length (bp)	GC no.	GC %	tm (°C)	tm-diff (°C)	product length (bp)
F: TGCCCTCTCTTGCAGAA(T->C)AAS	480-500	21	9	42.86	60	8	21 109
R: CTGCGTTCCCATTATGC	592-608	17	9	52.94	52		129

```
000001 TCTATTAGCC AGTTCCTAGC CATCCTCCTA ACTGTTGGAT CTGGAGGGTT AGCCTCAGAT
000061 CATTGCAGTT ATAAACAGTC CATCCCTGGC TATCTTTTCT AGAAGGGTCC AGTGTCATCT
000121 CAGCTAGGCA GAAACGTATT TCAACTCCCA AACTGTTTCA AGAGTTTGAA AAGCTCAGGC
000181 CCAGAAGAAA CAGAGTGAGC AGTTGAACAG TCTTACAACT GCTGCTTACA ACTTTATAAC
000241 AGGCCATAGG TTTAAAGTCC AAAAGGGCTC AAATTTCTAC CCCACGTTTG ACGACTCAGA
000301 CCTCAAGCCA GAGCTGAGCT GACTTCCCTG CGCCCAGGCC ACGGCGGTGA AATGAAGGCA
000361 CAGCAGCCCC GGCCAGGCCC TCCCAGTGGA GGCACAGGGG GACGGCGACA CCAACCACCC
000421 CAGCATCAGT AACCTGCACA CTCTTCTCCC TAGGTCTCCT GGAGGCAAGG CGACCTTGCT
000481 TGCCCTCTCT TGCAGAATAA SAAGGGGCTT AGCCACAGGA GTTGCTGGCA AGTGGAAAGA
000541 AGAACAAATG GTGAGCAGCA GTGCCGTTTA AGGACCAGTC GGGCTTGCAA ATGCATAATG
000601 GGAACGCAGA AAGGCCATGT GCAGATAGAC TGGACACAGC CGGAGGCAGC TCAAATCTGC
000661 CCCAACCTTC CTAGGGGCTA AAAGACGAGT GACAGAATCT GTAGCTGATG AGCTGGGGTG
000721 CCTCACGAAG GGTGGCCTTT TTACCTCTGT CCTTTGCTTT TTTCCTTTTA AAATCATAGC
000781 TGTAGTTGAT ATAAAATATA ACTTTGGGTG TGCAGCTTAA ATGATGGTGA GGACTGTCCT
000841 GCCTGCTGAA AGGCAGAAAC GCCTTTGCTT GTCATTGTTT CCCCAACATA GCATCCAAGA
000901 CAAAGGGGGT GCTCTGTGAC GCTTTGGAGC CTCTCTCACA TTCTTCTGAG CTGCTGGCAA
000961 ACCACAAGAT GAGAGTCCAG TGGCTCAAAG GGAAGCGAGT G
```

Fig. 2-26. The restriction enzymes and natural primer set information for rs45541837. The mutagenic position is underlined with green color in the F of primer pair field. Both forward and reverse primers are underlined with red color and the SNP is indicated by an arrow line. + is the sense strand and – is the antisense strand.

2.2.8 Example (8): TaqMan information providing

In above SNP (rs11540652), its TaqMan probe is not provided in both NCBI and SNP500Cancer (Fig. 2-15). However, some SNPs, such as rs1318703 which is shown on SNP ID example (2) (Fig. 2-27) and rs1202183 which is shown on SNP ID example (4) - ss1717832 (Fig. 2-28) provide TaqMan probe. When a SNP provide TaqMan probe, the icon of "TaqMan" will appear in the result of SNP-RFLP analysis.



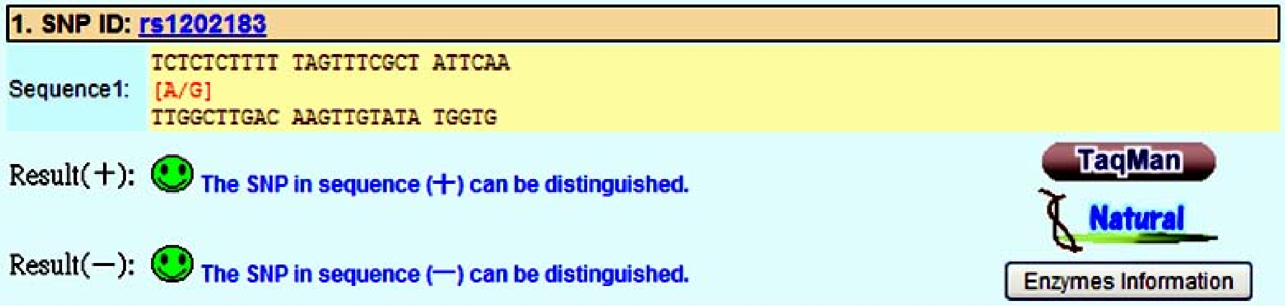
1. SNP ID: rs1318703

Sequence1: TATGAGGCAT ATTCGTTATT ATTTAC
[A/G]
TTTCTTTAAG TCATTTTTAT TATGT

Result(+): ☺ The SNP in sequence (+) can be distinguished.

Result(−): ☺ The SNP in sequence (−) can be distinguished.

Fig. 2-27. The icon of TaqMan appears in rs1318703 which is in SNP ID example (2).



1. SNP ID: rs1202183

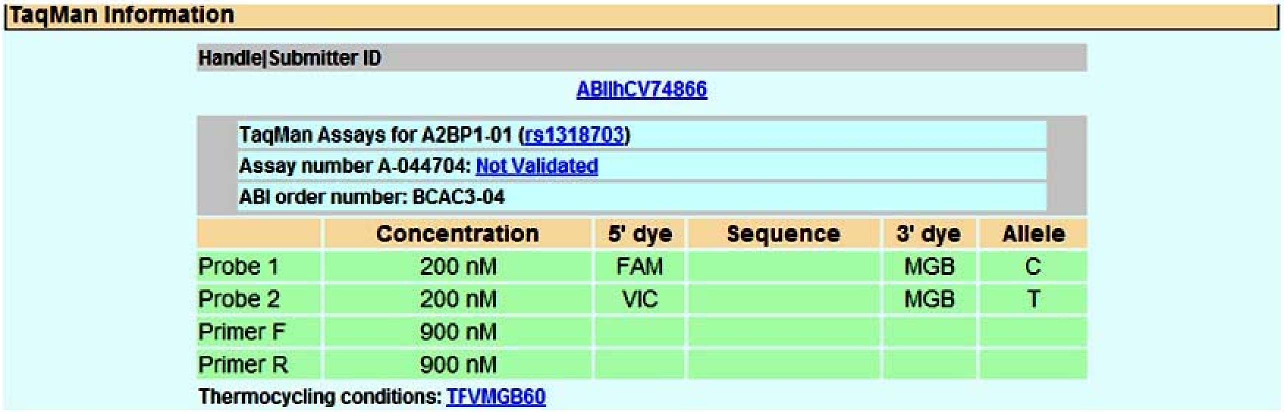
Sequence1: TCTCTCTTTT TAGTTTCGCT ATTCAA
[A/G]
TTGGCTTGAC AAGTTGTATA TGGTG

Result(+): ☺ The SNP in sequence (+) can be distinguished.

Result(−): ☺ The SNP in sequence (−) can be distinguished.

Fig. 2-28. The icon of TaqMan appears in rs1202183 which is in SNP ID example (4) - ss1717832.

User can click the icon of "TaqMan", and then TaqMan information for this SNP will be represented. Fig. 2-29 and Fig. 2-30 show the available TaqMan information for rs1318703 and rs1202183, respectively.



TaqMan Information

Handle|Submitter ID

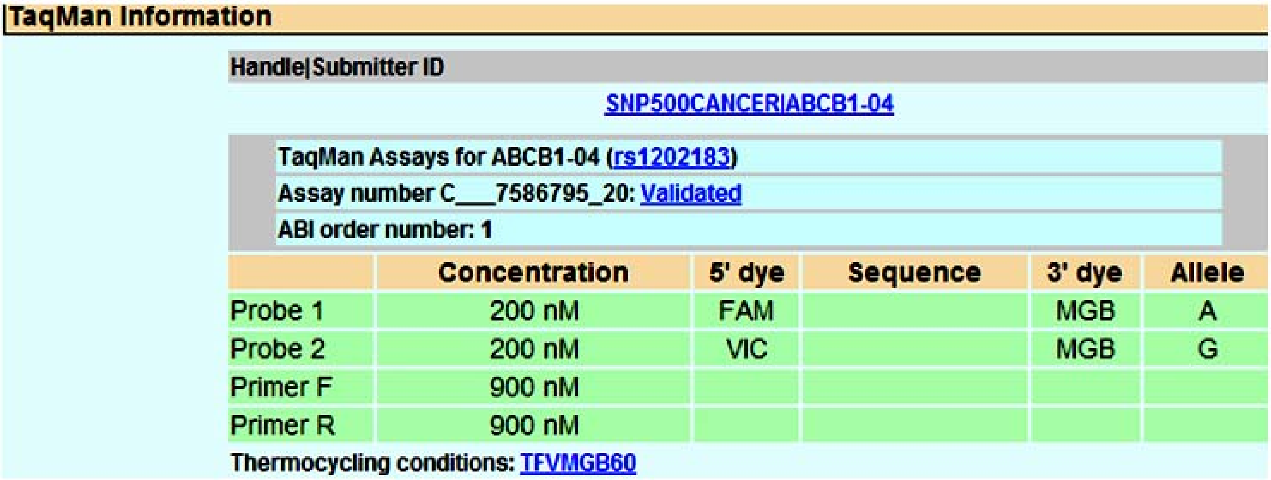
ABI|hCV74866

TaqMan Assays for A2BP1-01 (rs1318703)
Assay number A-044704: Not Validated
ABI order number: BCAC3-04

	Concentration	5' dye	Sequence	3' dye	Allele
Probe 1	200 nM	FAM		MGB	C
Probe 2	200 nM	VIC		MGB	T
Primer F	900 nM				
Primer R	900 nM				

Thermocycling conditions: TFVMGB60

Fig. 2-29. The TaqMan information for rs1318703 is shown after clicking the icon of "TaqMan" in Fig. 2-27.



TaqMan Information

Handle|Submitter ID

SNP500CANCER|ABCB1-04

TaqMan Assays for ABCB1-04 (rs1202183)

Assay number C___7586795_20: Validated

ABI order number: 1

	Concentration	5' dye	Sequence	3' dye	Allele
Probe 1	200 nM	FAM		MGB	A
Probe 2	200 nM	VIC		MGB	G
Primer F	900 nM				
Primer R	900 nM				

Thermocycling conditions: TFVMGB60

Fig. 2-30. The TaqMan information for rs1202183 is shown after clicking the icon of "TaqMan" in Fig. 2-28.

3. Function: SNP in fasta sequence input

3.1 Overview

SNP in fasta sequence format is acceptable to query the SNP-RFLP information in SNP-RFLPing 2. Single fasta sequence, multiple fasta sequences with SNPs in [dNTP1/dNTP2] or IUPAC formats, and special sequences with SNP in the indel (insertion and deletion), tri-allelic and tetra-allelic formats are all acceptable.

Fig. 3-1 shows a regular single fasta sequence.

Fig. 3-2 shows a single fasta sequence with line (press the enter key of keyboard) which is ignored in the system.

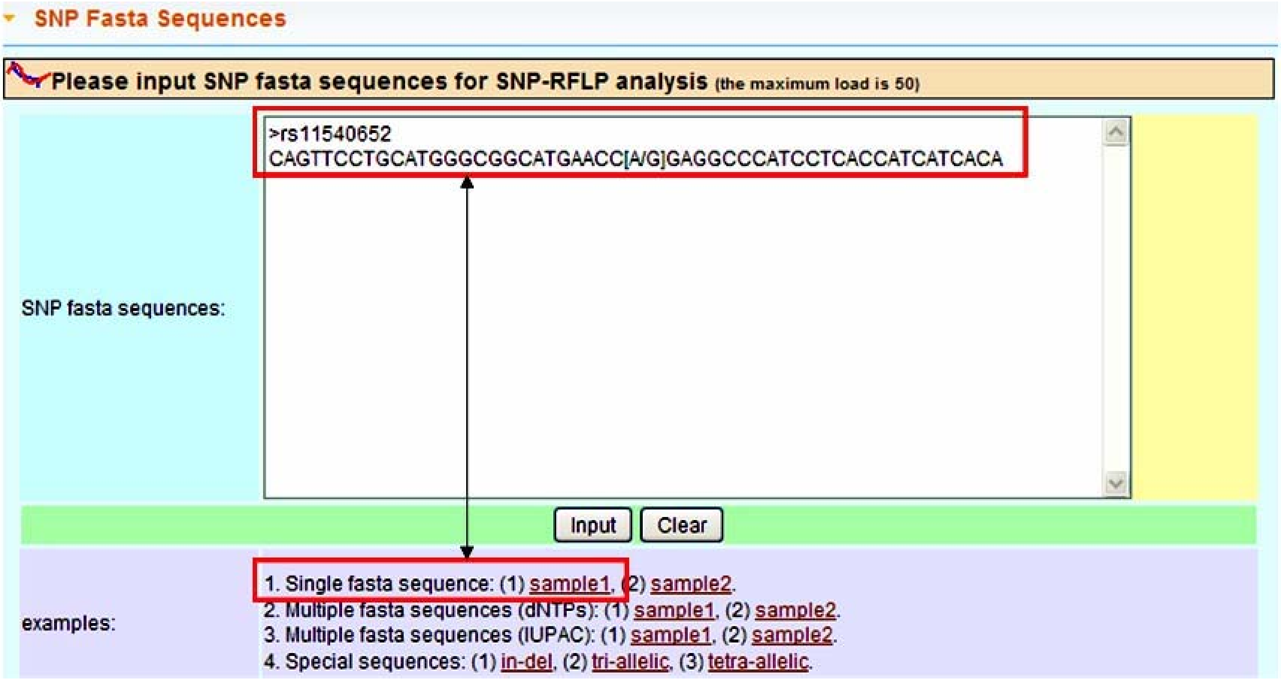
Fig. 3-3 and Fig. 3-5 show regular multiple fasta sequences with SNPs in [dNTP1/dNTP2] and IUPAC format, respectively.

Fig. 3-4 shows multiple fasta sequences with SNPs in [dNTP1/dNTP2].

Fig. 3-6 shows the IUPAC format with lines (press the enter key of keyboard) which are ignored.

Fig. 3-7 to Fig. 3-9 are special sequence with SNP in the in-del, tri-allelic and tetra-allelic formats, respectively. In the SNP fasta sequence input, the uppercase or lowercase letters are acceptable.

3.2 Example: Use SNP in fasta sequence format to analyze


SNP Fasta Sequences
Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 50)
Input
Clear

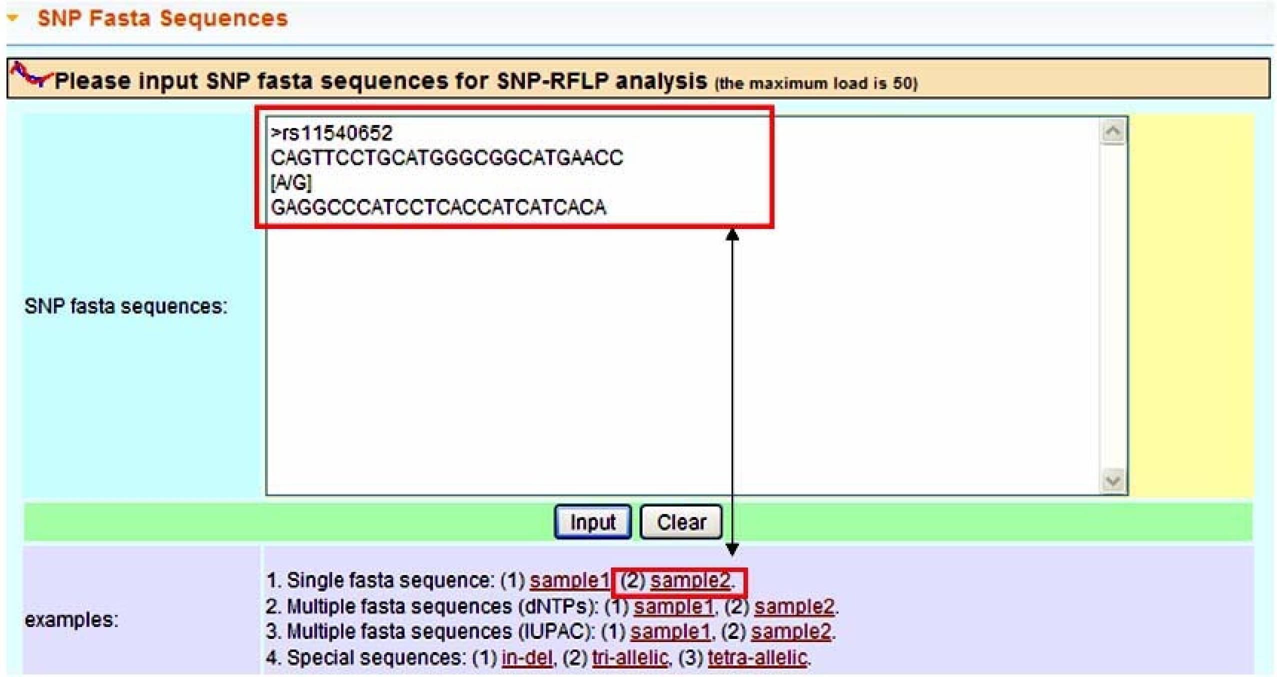
SNP fasta sequences:

```
>rs11540652
CAGTTCCTGCATGGGCGGCATGAACC[A/G]GAGGCCCATCCTCACCATCATCACA
```

examples:

1. Single fasta sequence: (1) sample1, (2) sample2.
2. Multiple fasta sequences (dNTPs): (1) sample1, (2) sample2.
3. Multiple fasta sequences (IUPAC): (1) sample1, (2) sample2.
4. Special sequences: (1) in-del, (2) tri-allelic, (3) tetra-allelic.

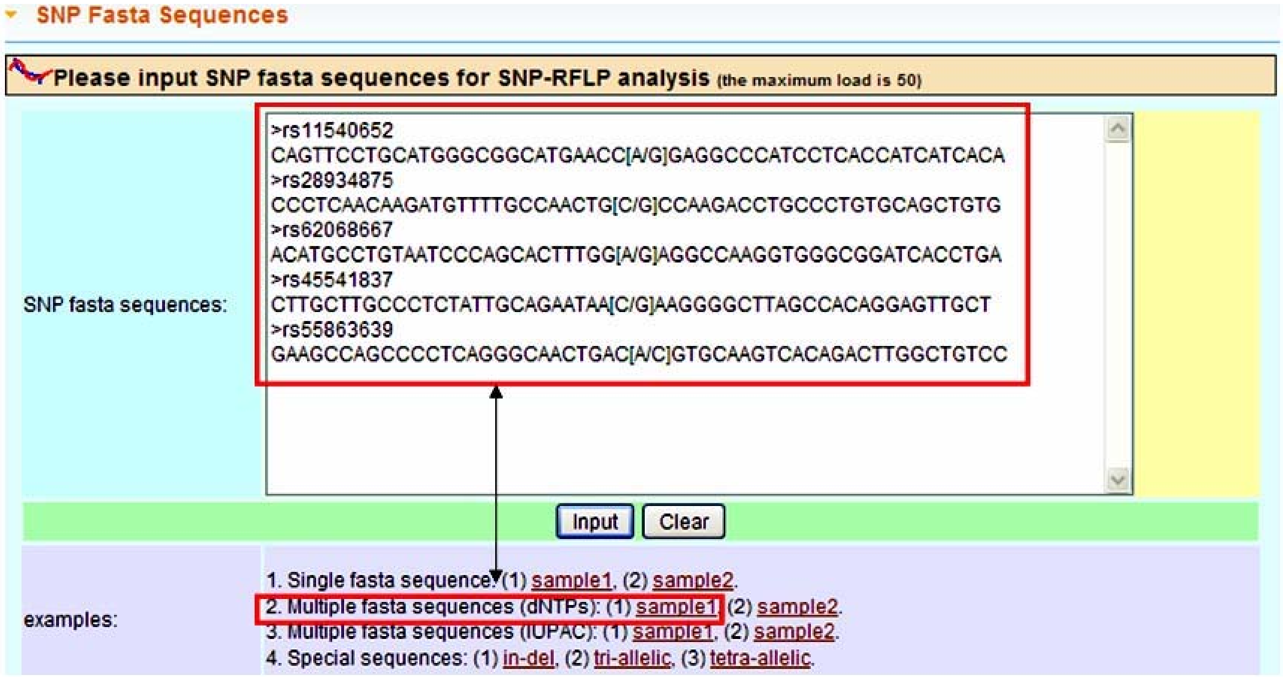
Fig. 3-1. SNP in a regular single fasta sequence.


SNP Fasta Sequences
Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 50)
SNP fasta sequences:
Input
Clear
examples:

```
>rs11540652
CAGTTCCTGCATGGGCGGCATGAACC
[A/G]
GAGGCCCATCCTCACCATCATCACA
```

1. Single fasta sequence: (1) sample1, (2) sample2.
2. Multiple fasta sequences (dNTPs): (1) sample1, (2) sample2.
3. Multiple fasta sequences (IUPAC): (1) sample1, (2) sample2.
4. Special sequences: (1) in-del, (2) tri-allelic, (3) tetra-allelic.

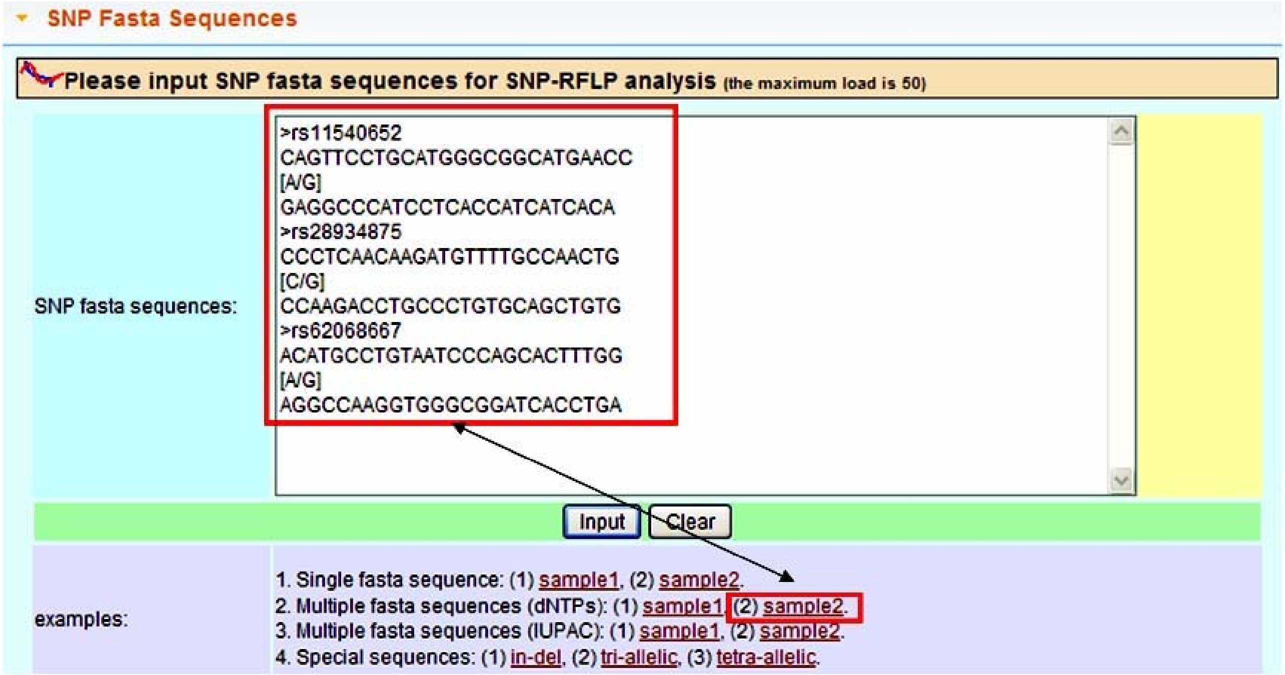
Fig. 3-2. SNP in a single fasta sequence with line (press the enter key of keyboard) which is ignored in the system. When the sequence is too long, the part of the sequence will list in the next line automatically. The system is able to regard them as the single sequence input.


SNP Fasta Sequences
Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 50)
SNP fasta sequences:
Input
Clear
examples:

```
>rs11540652
CAGTTCCTGCATGGGCGGCATGAACC[A/G]GAGGCCCATCCTCACCATCATCACA
>rs28934875
CCCTCAACAAGATGTTTTGCCAACTG[C/G]CCAAGACCTGCCCTGTGCAGCTGTG
>rs62068667
ACATGCCTGTAATCCCAGCACTTTGG[A/G]AGGCCAAGGTGGGCGGATCACCTGA
>rs45541837
CTTGCTTGCCCTCTATTGCAGAATAA[C/G]AAGGGGCTTAGCCACAGGAGTTGCT
>rs55863639
GAAGCCAGCCCCTCAGGGCAACTGAC[A/C]GTGCAAGTCACAGACTTGGCTGTCC
```

1. Single fasta sequence: (1) sample1, (2) sample2.
2. Multiple fasta sequences (dNTPs): (1) sample1, (2) sample2.
3. Multiple fasta sequences (IUPAC): (1) sample1, (2) sample2.
4. Special sequences: (1) in-del, (2) tri-allelic, (3) tetra-allelic.

Fig. 3-3. Regular multiple fasta sequences with SNPs in [dNTP1/dNTP2] format.



Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 50)

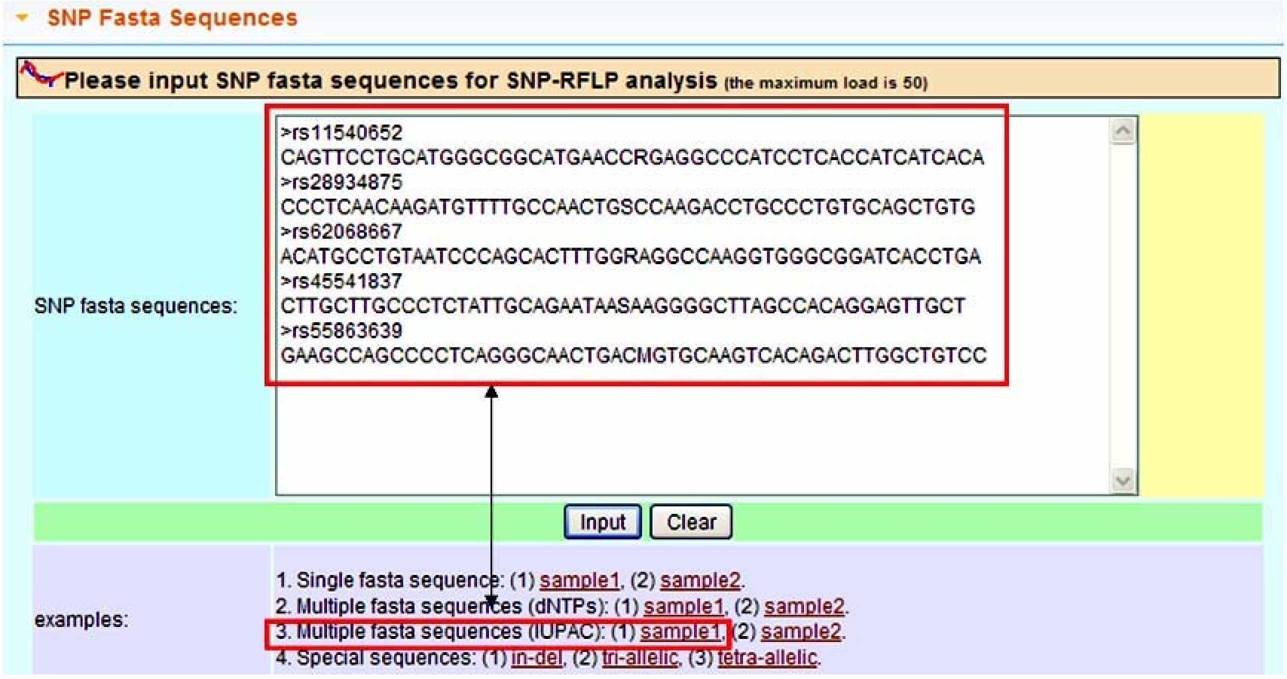
SNP fasta sequences:

```
>rs11540652
CAGTTCCTGCATGGGCGGCATGAACC
[A/G]
GAGGCCCATCCTCACCATCATCACA
>rs28934875
CCCTCAACAAGATGTTTTGCCAACTG
[C/G]
CCAAGACCTGCCCTGTGCAGCTGTG
>rs62068667
ACATGCCTGTAATCCCAGCACTTTGG
[A/G]
AGGCCAAGGTGGGCGGATCACCTGA
```

examples:

1. Single fasta sequence: (1) sample1, (2) sample2.
2. Multiple fasta sequences (dNTPs): (1) sample1, (2) sample2.
3. Multiple fasta sequences (IUPAC): (1) sample1, (2) sample2.
4. Special sequences: (1) in-del, (2) tri-allelic, (3) tetra-allelic.

Fig. 3-4. Multiple fasta sequences with SNPs in [dNTP1/dNTP2] format with line (press the enter key of keyboard) which are ignored in the system. When the sequence is too long, the part of the sequence will list in the next line automatically. The system is able to regard them as the single sequence input.



Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 50)

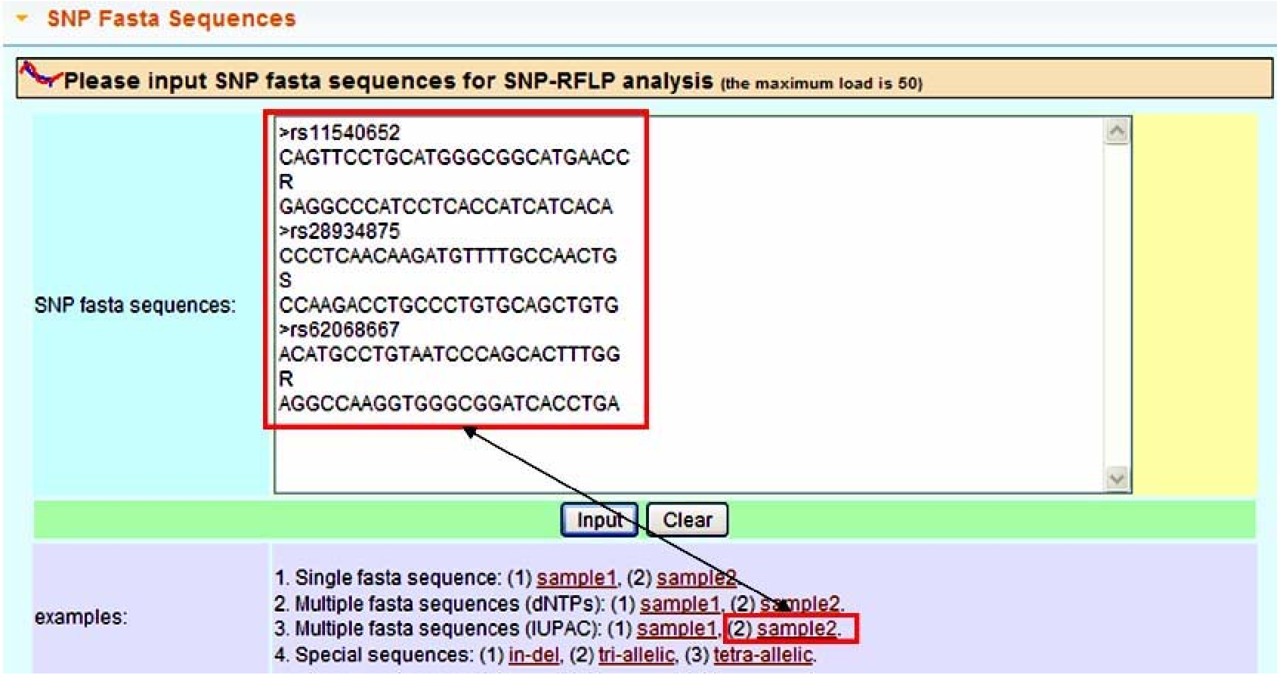
SNP fasta sequences:

```
>rs11540652
CAGTTCCTGCATGGGCGGCATGAACCRGAGGCCCATCCTCACCATCATCACA
>rs28934875
CCCTCAACAAGATGTTTTGCCAACTGSCCAAGACCTGCCCTGTGCAGCTGTG
>rs62068667
ACATGCCTGTAATCCCAGCACTTTGGRAGGCCAAGGTGGGCGGATCACCTGA
>rs45541837
CTTGCTTGCCCTCTATTGCAGAATAASAAGGGGCTTAGCCACAGGAGTTGCT
>rs55863639
GAAGCCAGCCCCTCAGGGCAACTGACMGTGCAAGTCACAGACTTGGCTGTCC
```

examples:

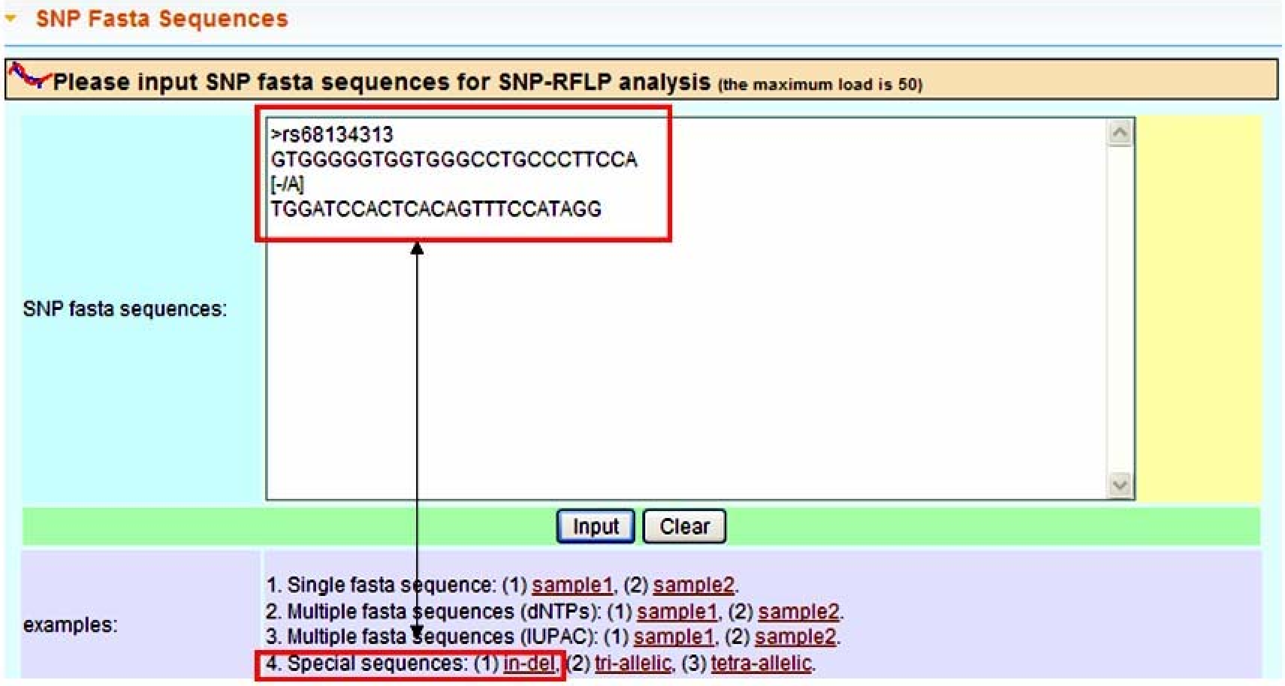
1. Single fasta sequence: (1) sample1, (2) sample2.
2. Multiple fasta sequences (dNTPs): (1) sample1, (2) sample2.
3. Multiple fasta sequences (IUPAC): (1) sample1, (2) sample2.
4. Special sequences: (1) in-del, (2) tri-allelic, (3) tetra-allelic.

Fig. 3-5. Regular multiple fasta sequences with SNPs in IUPAC format.


SNP Fasta Sequences
Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 50)
SNP fasta sequences:
Input
Clear
examples:
1. Single fasta sequence: (1) sample1, (2) sample2.
2. Multiple fasta sequences (dNTPs): (1) sample1, (2) sample2.
3. Multiple fasta sequences (IUPAC): (1) sample1, (2) sample2.
4. Special sequences: (1) in-del, (2) tri-allelic, (3) tetra-allelic.

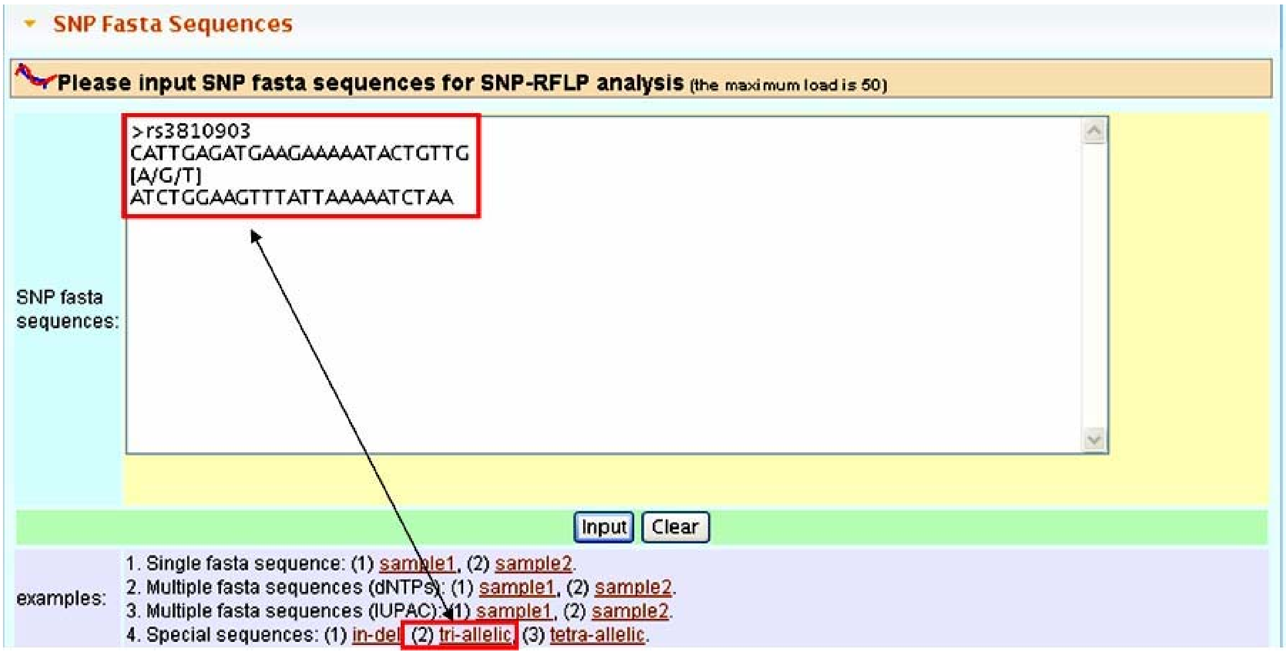
```
>rs11540652
CAGTTCCTGCATGGGCGGCATGAACC
R
GAGGCCCATCCTCACCATCATCACA
>rs28934875
CCCTCAACAAGATGTTTTGCCAACTG
S
CCAAGACCTGCCCTGTGCAGCTGTG
>rs62068667
ACATGCCTGTAATCCCAGCACTTTGG
R
AGGCCAAGGTGGGCGGATCACCTGA
```

Fig. 3-6. Multiple fasta sequences with SNPs in IUPAC format with line (press the enter key of keyboard) which is ignored in the system. R = G or A; S = G or C. When the sequence is too long, the part of the sequence will list in the next line automatically. The system is able to regard them as the single sequence input.


SNP Fasta Sequences
Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 50)
SNP fasta sequences:
Input
Clear
examples:
1. Single fasta sequence: (1) sample1, (2) sample2.
2. Multiple fasta sequences (dNTPs): (1) sample1, (2) sample2.
3. Multiple fasta sequences (IUPAC): (1) sample1, (2) sample2.
4. Special sequences: (1) in-del, (2) tri-allelic, (3) tetra-allelic.

```
>rs68134313
GTGGGGGTGGTGGGCCTGCCCTTCCA
[-/A]
TGGATCCACTCACAGTTTCCATAGG
```

Fig. 3-7. Special sequence with SNP in the in-del format. When the sequence is too long, the part of the sequence will list in the next line automatically. The system is able to regard them as the single sequence input.



SNP Fasta Sequences

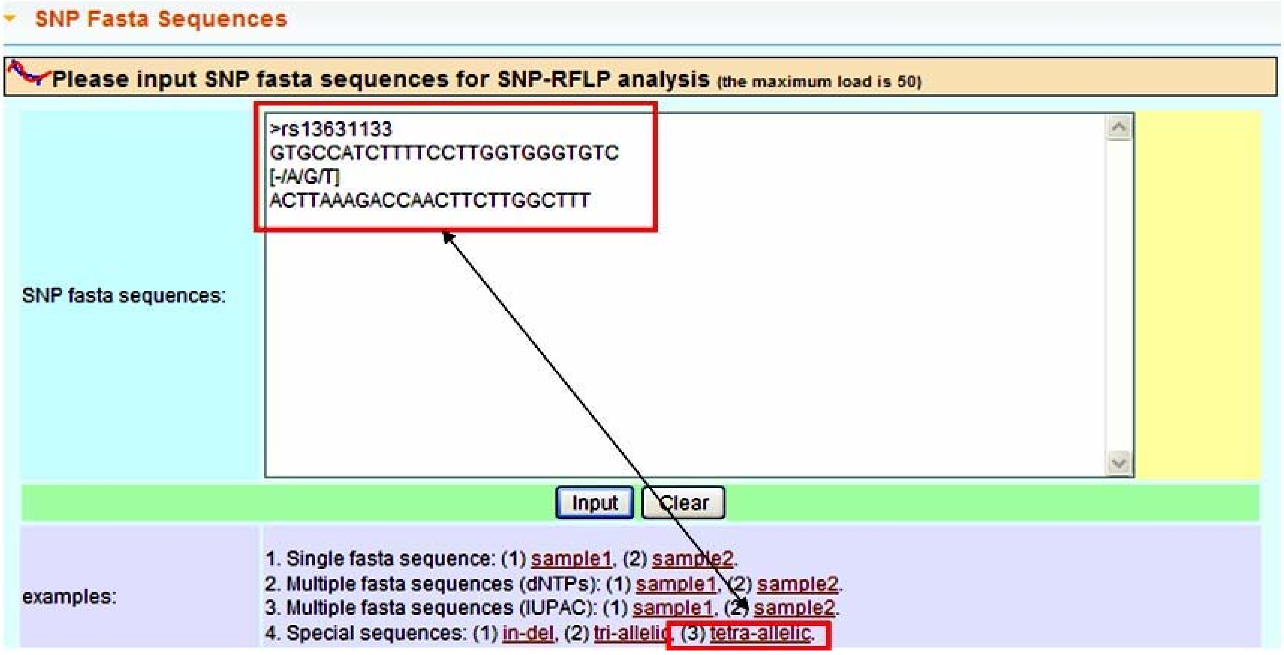
Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 50)

SNP fasta sequences:

```
>rs3810903
CATTGAGATGAAGAAAAATACTGTTG
[A/G/T]
ATCTGGAAGTTTATTAAAAATCTAA
```

examples:
1. Single fasta sequence: (1) sample1, (2) sample2.
2. Multiple fasta sequences (dNTPs): (1) sample1, (2) sample2.
3. Multiple fasta sequences (IUPAC): (1) sample1, (2) sample2.
4. Special sequences: (1) in-del, (2) tri-allelic, (3) tetra-allelic.

Fig. 3-8. Special sequence with SNP in the tri-allelic format. When the sequence is too long, the part of the sequence will list in the next line automatically. The system is able to regard them as the single sequence input.



SNP Fasta Sequences

Please input SNP fasta sequences for SNP-RFLP analysis (the maximum load is 50)

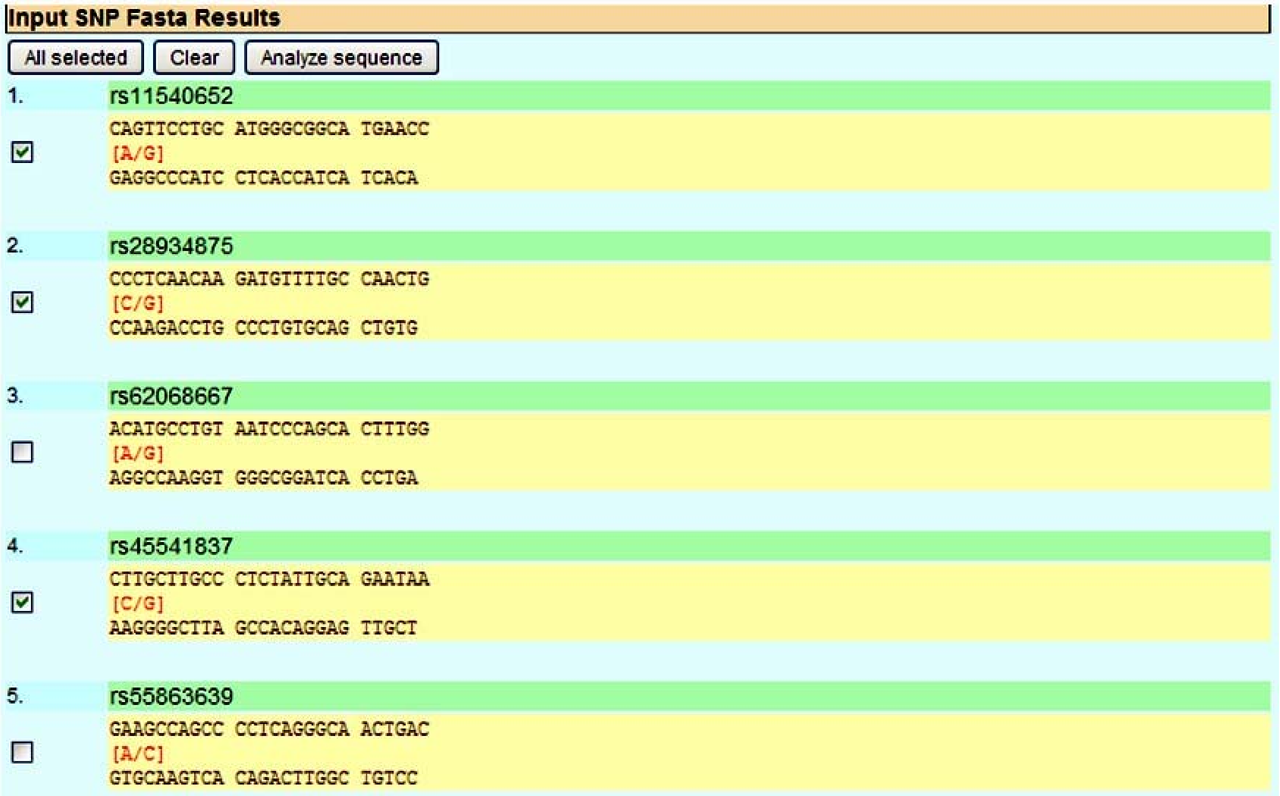
SNP fasta sequences:

```
>rs13631133
GTGCCATCTTTTCCTTGGTGGGTGTC
[-/A/G/T]
ACTTAAAGACCAACTTCTTGGCTTT
```

examples:
1. Single fasta sequence: (1) sample1, (2) sample2.
2. Multiple fasta sequences (dNTPs): (1) sample1, (2) sample2.
3. Multiple fasta sequences (IUPAC): (1) sample1, (2) sample2.
4. Special sequences: (1) in-del, (2) tri-allelic, (3) tetra-allelic.

Fig. 3-9. Special sequence with SNP in the tetra-allelic format. When the sequence is too long, the part of the sequence will list in the next line automatically. The system is able to regard them as the single sequence input.

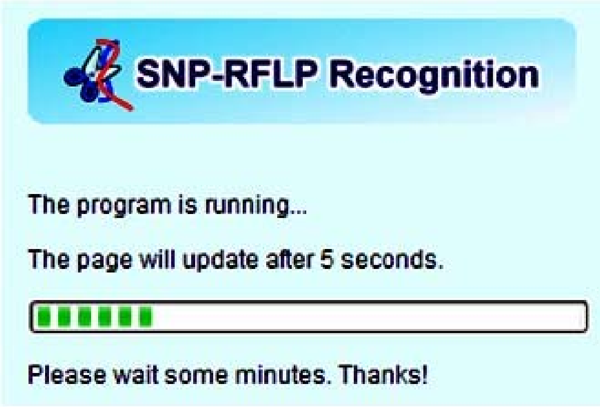
After inputting the SNP fasta sequence as mentioned above, users can click the “Input” button, and all information which is not belong to ‘A’, ‘T’, ‘C’, ‘T’, [dNTP1/dNTP2/…/ dNTPn], or IUPAC will be filtered out. The results for SNP fasta sequence input are shown as Fig. 3-10, including: (1) “All selected”, (2) “Clear”, (3) “Analyze sequence”, (4) the description for sequence, and (5) SNP sequence. Users can check the desired SNPs and click the “Analyze sequence” to next step. If the all SNPs need be analyzed, the “All selected” button provides convenience to check all SNPs. The “Clear” button provides cancel all SNP checked.



```
1.     rs11540652
☑      CAGTTCCTGC ATGGGCGGCA TGAACC
       [A/G]
       GAGGCCCATC CTCACCATCA TCACA

2.     rs28934875
☑      CCCTCAACAA GATGTTTTGC CAACTG
       [C/G]
       CCAAGACCTG CCCTGTGCAG CTGTG

3.     rs62068667
☐      ACATGCCTGT AATCCCAGCA CTTTGG
       [A/G]
       AGGCCAAGGT GGGCGGATCA CCTGA

4.     rs45541837
☑      CTTGCTTGCC CTCTATTGCA GAATAA
       [C/G]
       AAGGGGCTTA GCCACAGGAG TTGCT

5.     rs55863639
☐      GAAGCCAGCC CCTCAGGGCA ACTGAC
       [A/C]
       GTGCAAGTCA CAGACTTGGC TGTCC
```

Fig. 3-10. The results for SNP fasta sequence input.

When clicking the “Analyze sequence”, users need to wait for the analysis of SNP-RFLP recognition (Fig. 3-11). Then, the results for SNP-RFLP recognition will be presented (Fig. 3-12). Finally, users can select the interested SNPs as described above.

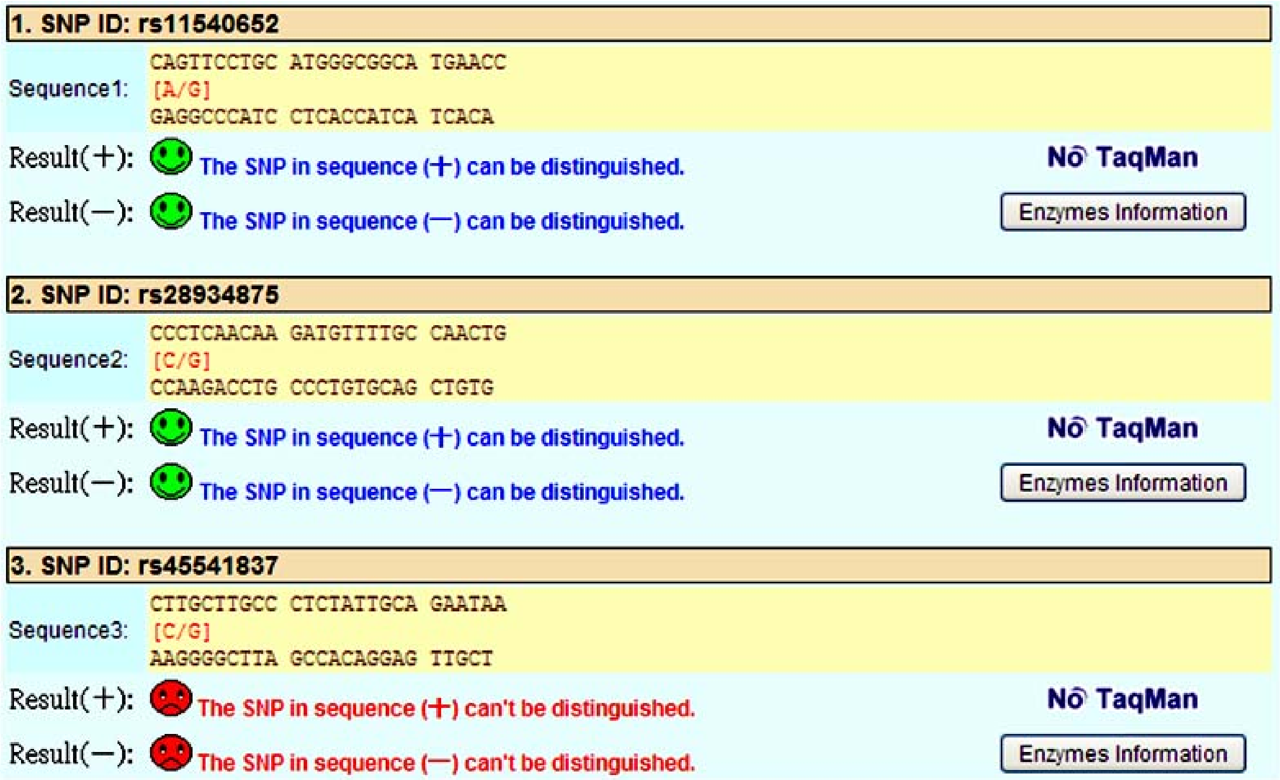


The program is running...

The page will update after 5 seconds.

Please wait some minutes. Thanks!

Fig. 3-11. Wait for SNP-RFLP recognition.



1. SNP ID: rs11540652

Sequence1: CAGTTCCTGC ATGGGCGGCA TGAACC [A/G] GAGGCCCATC CTCACCATCA TCACA

Result(+): ☺ The SNP in sequence (+) can be distinguished.

Result(−): ☺ The SNP in sequence (−) can be distinguished.

2. SNP ID: rs28934875

Sequence2: CCCTCAACAA GATGTTTTGC CAACTG [C/G] CCAAGACCTG CCCTGTGCAG CTGTG

Result(+): ☺ The SNP in sequence (+) can be distinguished.

Result(−): ☺ The SNP in sequence (−) can be distinguished.

3. SNP ID: rs45541837

Sequence3: CTTGCTTGCC CTCTATTGCA GAATAA [C/G] AAGGGGCTTA GCCACAGGAG TTGCT

Result(+): ☹ The SNP in sequence (+) can't be distinguished.

Result(−): ☹ The SNP in sequence (−) can't be distinguished.

Fig. 3-12. The results for SNP-RFLP recognition.

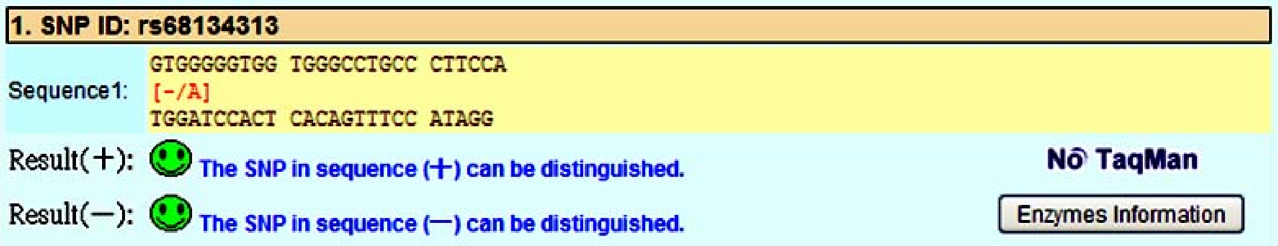
When uses input a fasta sequence with SNP in the in-del (insertion and deletion), tri-allelic, and tetra-allelic formats, their results are described as follows.

For SNP in the in-del format, the results are similar to di-allelic SNP (Fig. 3-13 and Fig. 3-14).

For SNP in the tri-allelic format, the results for restriction enzymes information will present three sequence IDs with their corresponding available restriction enzymes for three different alleles (Fig. 3-15).

For SNP in the tetra-allelic format, the results for restriction enzymes information will present four sequence IDs with their corresponding available restriction enzymes for four different alleles (Fig. 3-16).

For input SNP in fasta sequence format, the results for SNP-RFLP recognition will show the natural or mutagenic design results, because the SNP flanking sequence is not provided by the user input sequence.



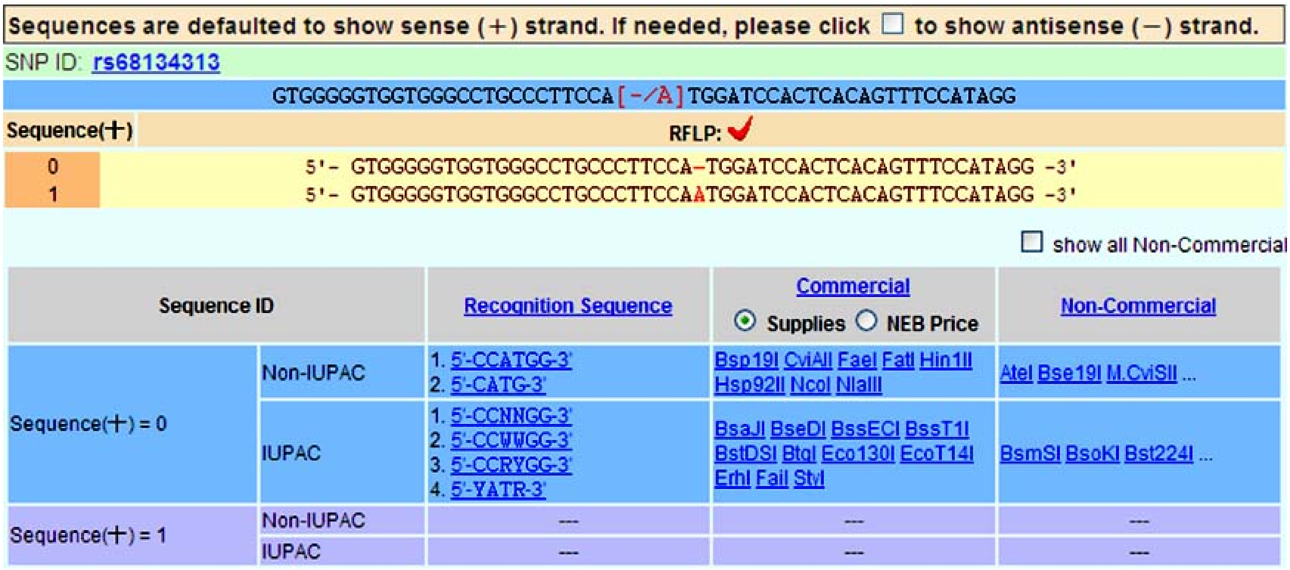
1. SNP ID: rs68134313

Sequence1: GTGGGGGTGG TGGGCCTGCC CTTCCA [-/A] TGGATCCACT CACAGTTTCC ATAGG

Result(+): ☺ The SNP in sequence (+) can be distinguished.

Result(−): ☺ The SNP in sequence (−) can be distinguished.

Fig. 3-13. The result for SNP-RFLP recognition with SNP in the in-del farmat.



Sequences are defaulted to show sense (+) strand. If needed, please click ☐ to show antisense (−) strand.

SNP ID: rs68134313

GTGGGGGTGGTGGGCCTGCCCTTCCA[-/A]TGGATCCACTCACAGTTTCCATAGG

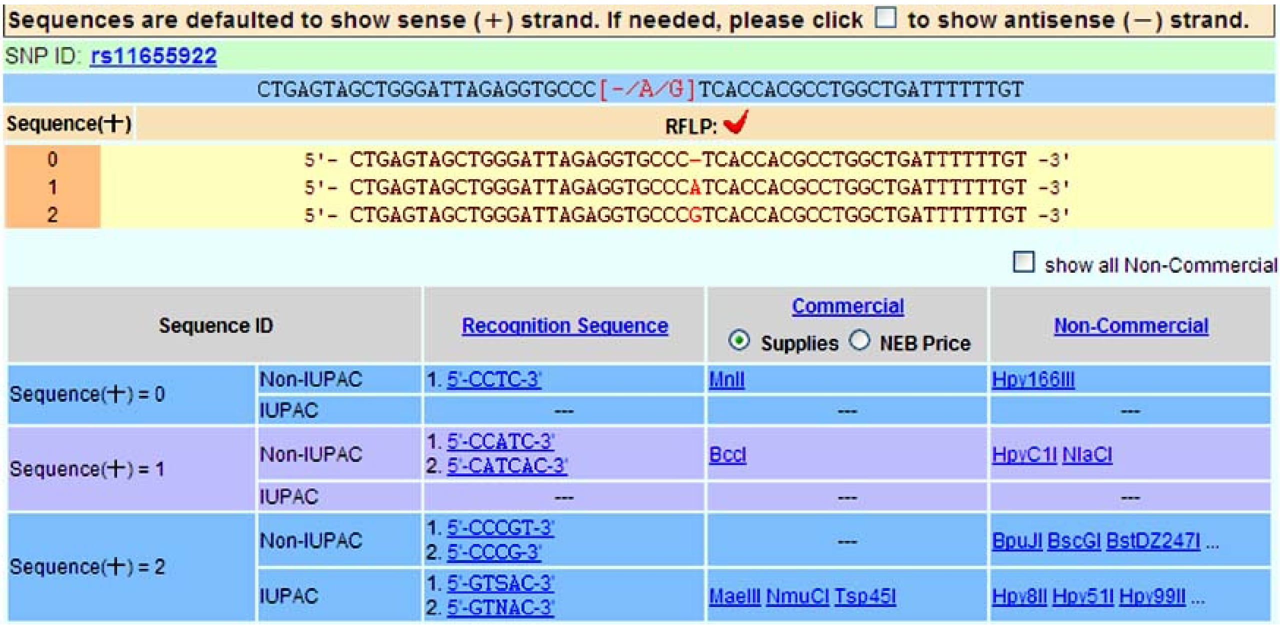
Sequence(+) RFLP: ✓

0 5'- GTGGGGGTGGTGGGCCTGCCCTTCCA-TGGATCCACTCACAGTTTCCATAGG -3'
1 5'- GTGGGGGTGGTGGGCCTGCCCTTCCAATGGATCCACTCACAGTTTCCATAGG -3'

☐ show all Non-Commercial

Sequence ID		Recognition Sequence	Commercial ● Supplies ○ NEB Price	Non-Commercial
Sequence(+) = 0	Non-IUPAC	1. 5'-CCATGG-3' 2. 5'-CATG-3'	Bsp19I CviAII FaeI FatI Hin1II Hsp92II NcoI NlaIII	AteI Bse19I M.CviSII ...
	IUPAC	1. 5'-CCNNGG-3' 2. 5'-CCWWGG-3' 3. 5'-CCRYGG-3' 4. 5'-YATR-3'	BsaJI BseDI BssECI BssT1I BstDSI BtgI Eco130I EcoT14I ErhI FaiI StyI	BsmSI BsoKI Bst224I ...
Sequence(+) = 1	Non-IUPAC	---	---	---
	IUPAC	---	---	---

Fig. 3-14. The available restriction enzyme information for the sense strand with SNP in the in-del format.



Sequences are defaulted to show sense (+) strand. If needed, please click ☐ to show antisense (−) strand.

SNP ID: rs11655922

CTGAGTAGCTGGGATTAGAGGTGCCC[-/A/G]TCACCACGCCTGGCTGATTTTTTGT

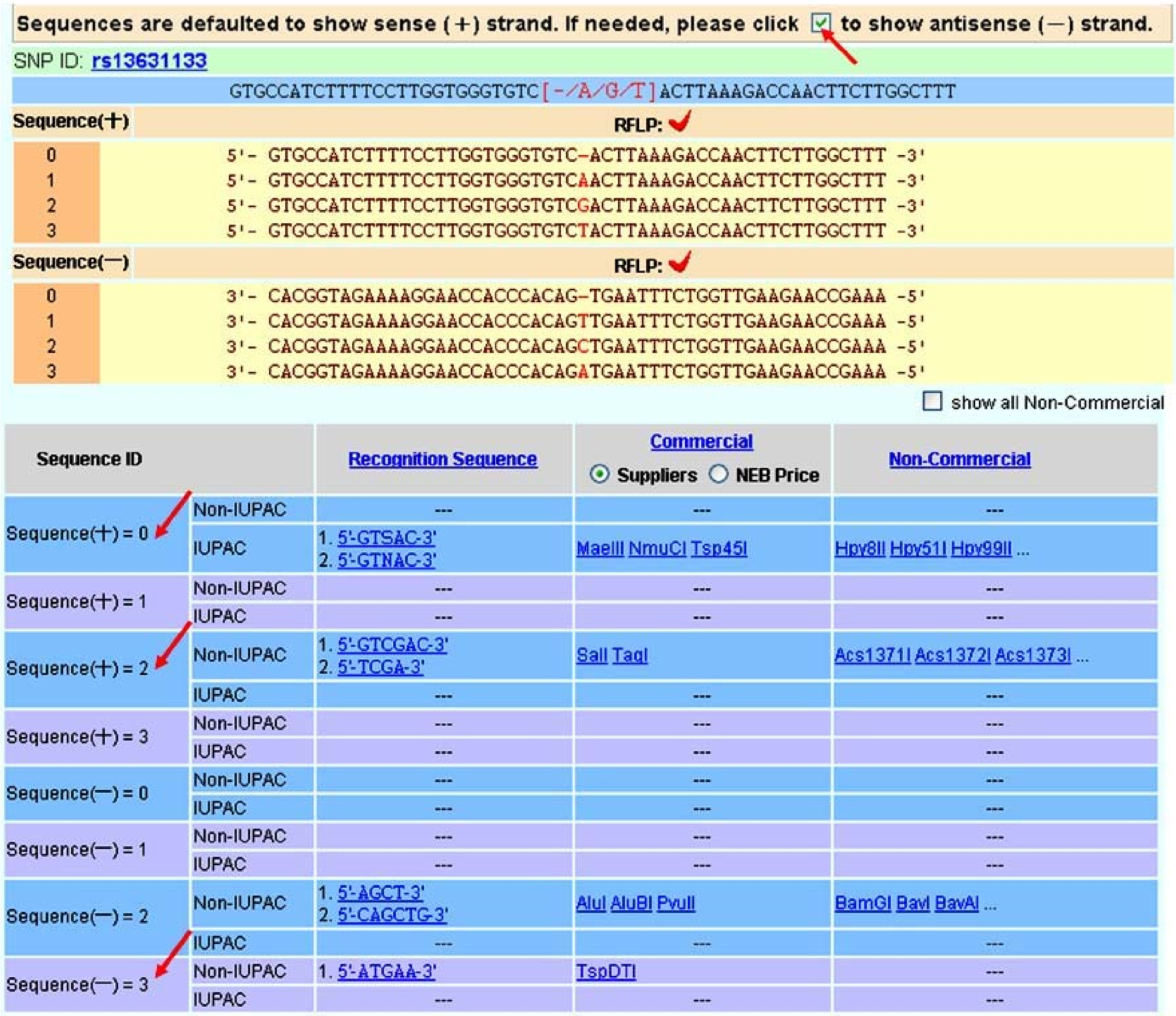
Sequence(+) RFLP: ✓

0 5'- CTGAGTAGCTGGGATTAGAGGTGCCC-TCACCACGCCTGGCTGATTTTTTGT -3'
1 5'- CTGAGTAGCTGGGATTAGAGGTGCCCATCACCACGCCTGGCTGATTTTTTGT -3'
2 5'- CTGAGTAGCTGGGATTAGAGGTGCCCGTCACCACGCCTGGCTGATTTTTTGT -3'

☐ show all Non-Commercial

Sequence ID		Recognition Sequence	Commercial ● Supplies ○ NEB Price	Non-Commercial
Sequence(+) = 0	Non-IUPAC	1. 5'-CCTC-3'	MnlI	Hpy166III
	IUPAC	---	---	---
Sequence(+) = 1	Non-IUPAC	1. 5'-CCATC-3' 2. 5'-CATCAC-3'	BccI	HpyC1I NlaCI
	IUPAC	---	---	---
Sequence(+) = 2	Non-IUPAC	1. 5'-CCCGT-3' 2. 5'-CCCG-3'	---	BpuJI BscGI BstDZ247I ...
	IUPAC	1. 5'-GTSAC-3' 2. 5'-GTNAC-3'	MaeIII NmuCI Tsp45I	Hpy8II Hpy51I Hpy99II ...

Fig. 3-15. The available restriction enzyme information for the sense strand with SNP in the tri-allelic format.



Sequences are defaulted to show sense (+) strand. If needed, please click ☑ to show antisense (—) strand.

SNP ID: rs13631133

GTGCCATCTTTTCCTTGGTGGGTGTC[-/A/G/T]ACTTAAAGACCAACTTCTTGGCTTT

Sequence(+) RFLP: ✓

```
0   5'- GTGCCATCTTTTCCTTGGTGGGTGTC-ACTTAAAGACCAACTTCTTGGCTTT -3'
1   5'- GTGCCATCTTTTCCTTGGTGGGTGTCAACTTAAAGACCAACTTCTTGGCTTT -3'
2   5'- GTGCCATCTTTTCCTTGGTGGGTGTCGACTTAAAGACCAACTTCTTGGCTTT -3'
3   5'- GTGCCATCTTTTCCTTGGTGGGTGTCTACTTAAAGACCAACTTCTTGGCTTT -3'
```

Sequence(—) RFLP: ✓

```
0   3'- CACGGTAGAAAAGGAACCACCCACAG-TGAATTTCTGGTTGAAGAACCGAAA -5'
1   3'- CACGGTAGAAAAGGAACCACCCACAGTTGAATTTCTGGTTGAAGAACCGAAA -5'
2   3'- CACGGTAGAAAAGGAACCACCCACAGCTGAATTTCTGGTTGAAGAACCGAAA -5'
3   3'- CACGGTAGAAAAGGAACCACCCACAGATGAATTTCTGGTTGAAGAACCGAAA -5'
```

☐ show all Non-Commercial

Sequence ID		Recognition Sequence	Commercial ● Suppliers ○ NEB Price	Non-Commercial
Sequence(+) = 0	Non-IUPAC	---	---	---
	IUPAC	1. 5'-GTSAC-3' 2. 5'-GTNAC-3'	MaeIII NmuCI Tsp45I	Hpy8II Hpy51I Hpy99II ...
Sequence(+) = 1	Non-IUPAC	---	---	---
	IUPAC	---	---	---
Sequence(+) = 2	Non-IUPAC	1. 5'-GTCGAC-3' 2. 5'-TCGA-3'	SalI TaqI	Acs1371I Acs1372I Acs1373I ...
	IUPAC	---	---	---
Sequence(+) = 3	Non-IUPAC	---	---	---
	IUPAC	---	---	---
Sequence(—) = 0	Non-IUPAC	---	---	---
	IUPAC	---	---	---
Sequence(—) = 1	Non-IUPAC	---	---	---
	IUPAC	---	---	---
Sequence(—) = 2	Non-IUPAC	1. 5'-AGCT-3' 2. 5'-CAGCTG-3'	AluI AluBI PvuII	BamGI BavI BavAI ...
	IUPAC	---	---	---
Sequence(—) = 3	Non-IUPAC	1. 5'-ATGAA-3'	TspDTI	---
	IUPAC	---	---	---

Fig. 3-16. The available restriction enzyme information for the sense strand with SNP in the tetra-allelic format. In this case, the enzymes for the antisense strand are chosen (*as indicated by red arrow in the top*). In the left side, three red arrow lines indicate that three nucleotides are distinguished between each other. The sequence 1 is unable to find the suitable enzyme but it is still distinct to others because it is unable to cut by the enzymes listed for sequences 0, 2, and 3.

4. Function: Multiple SNPs within one sequence

4.1 Overview

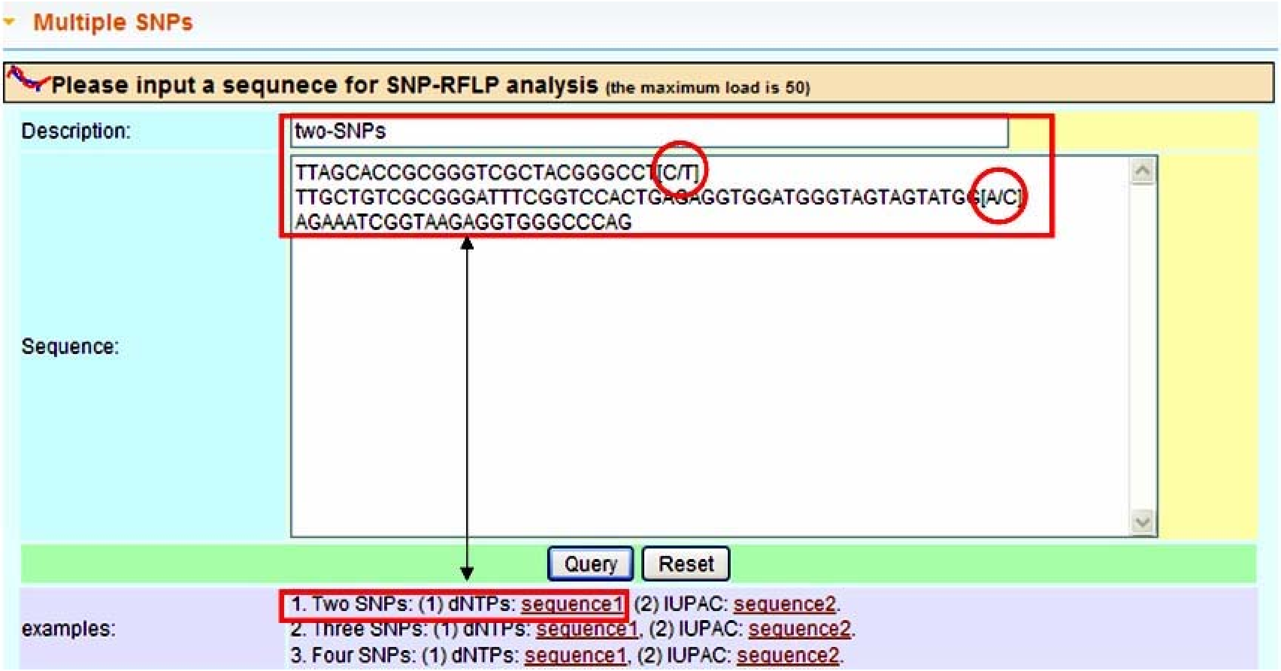
In the function for multiple SNPs within one sequence, up to 50 SNPs representing in the [dNTP1/dNTP2] or IUPAC formats within the input sequence are acceptable for analysis. The flanking sequences for two nearby SNPs should not be overlapped within 6 nucleotides. There are two input areas in the function: one is “Description” which is a title for describing input sequence and the other is “Sequence” which is input multiple SNPs sequence.

Fig. 4-1 and Fig. 4-2 are two SNPs within one sequence in the [dNTP1/dNTP2] and IUPAC formats, respectively.

Fig. 4-3 and Fig. 4-4 are three SNPs within one sequence in the [dNTP1/dNTP2] and IUPAC formats, respectively.

Fig. 4-5 and Fig. 4-6 are four SNPs within one sequence in the [dNTP1/dNTP2] and IUPAC formats, respectively.

4.2 Example: Multiple SNPs within one sequence input


Multiple SNPs
Please input a sequnece for SNP-RFLP analysis (the maximum load is 50)
Query
Reset

Description: two-SNPs

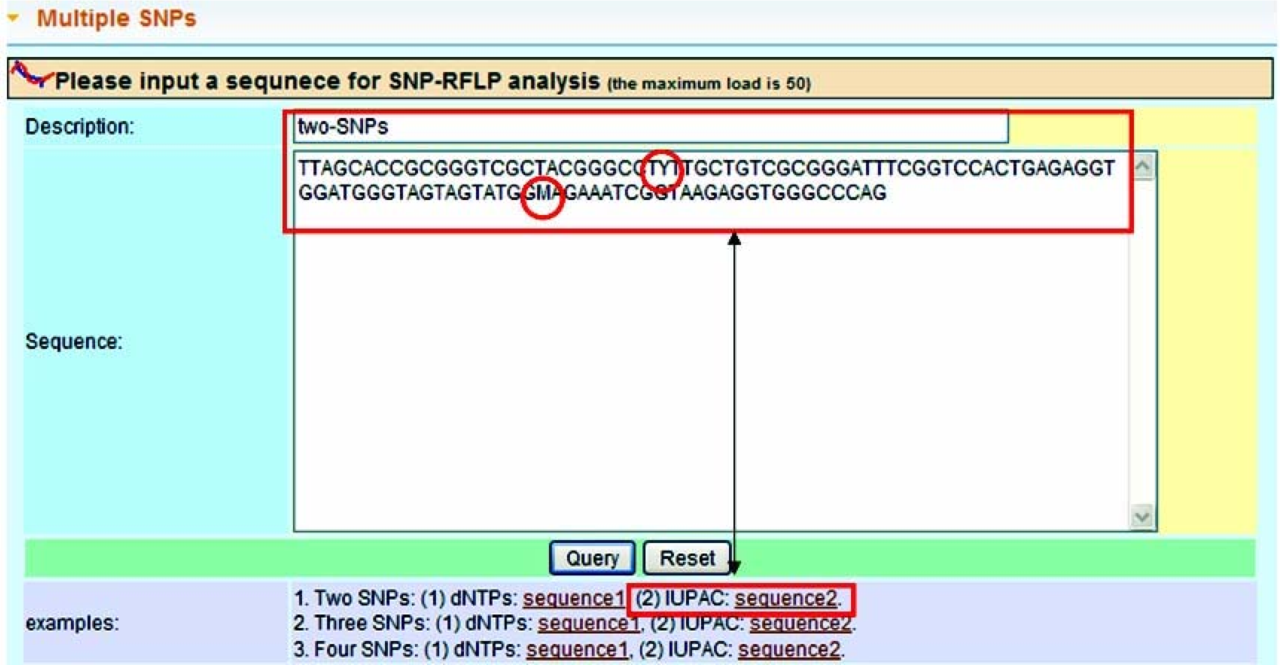
Sequence:

```
TTAGCACCGCGGGTCGCTACGGGCCT[C/T]
TTGCTGTCGCGGGATTTCGGTCCACTGAGAGGTGGATGGGTAGTAGTATGG[A/C]
AGAAATCGGTAAGAGGTGGGCCCAG
```

examples:

1. Two SNPs: (1) dNTPs: sequence1 (2) IUPAC: sequence2.
2. Three SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
3. Four SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.

Fig. 4-1. Two SNPs within one sequence in the [dNTP1/dNTP2] format. When the sequence is too long, the part of the sequence will list in the next line automatically. The system is able to regard them as the single sequence input. If the SNP is marked in IUPAC code, the sequence is continued as shown in Fig. 4-2.



Multiple SNPs
Please input a sequnece for SNP-RFLP analysis (the maximum load is 50)
Query
Reset

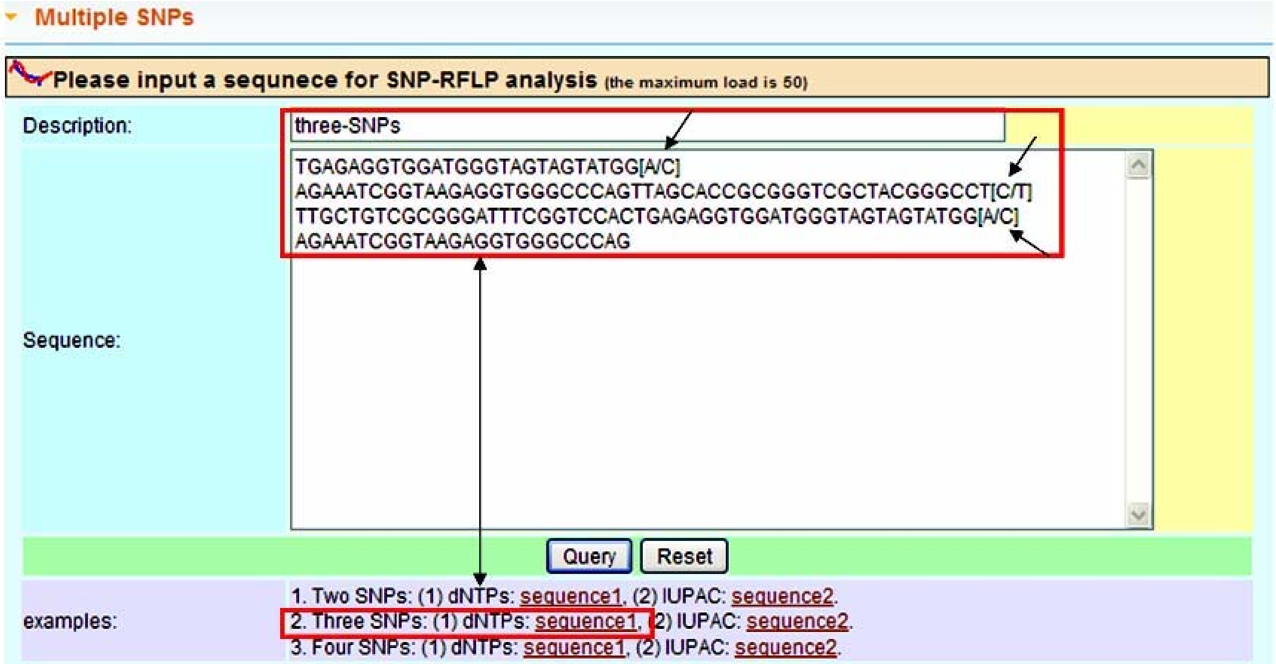
Description: two-SNPs

Sequence:

TTAGCACCGCGGGTCGCTACGGGCCTYTTGCTGTCGCGGGATTTCGGTCCACTGAGAGGT
GGATGGGTAGTAGTATGGMAGAAATCGGTAAGAGGTGGGCCCAG

examples:
1. Two SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
2. Three SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
3. Four SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.

Fig. 4-2. Two SNPs within one sequence in the IUPAC format.



Multiple SNPs
Please input a sequnece for SNP-RFLP analysis (the maximum load is 50)
Query
Reset

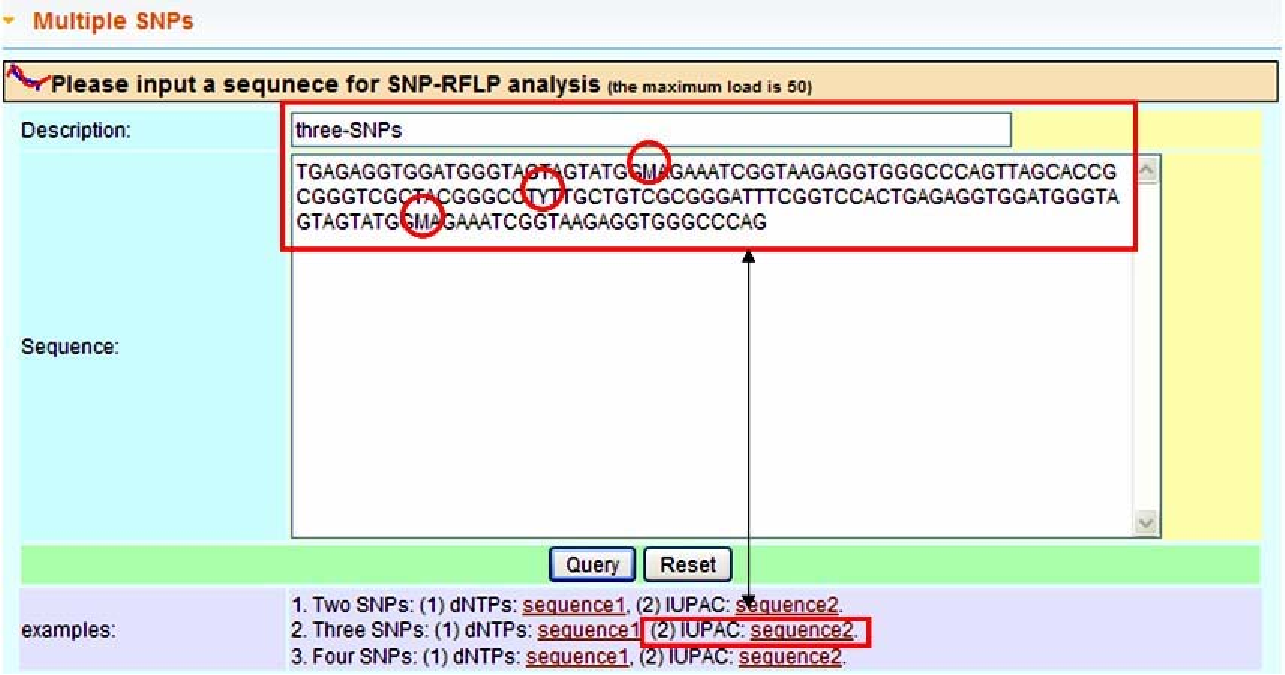
Description: three-SNPs

Sequence:

TGAGAGGTGGATGGGTAGTAGTATGG[A/C]
AGAAATCGGTAAGAGGTGGGCCCAGTTAGCACCGCGGGTCGCTACGGGCCT[C/T]
TTGCTGTCGCGGGATTTCGGTCCACTGAGAGGTGGATGGGTAGTAGTATGG[A/C]
AGAAATCGGTAAGAGGTGGGCCCAG

examples:
1. Two SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
2. Three SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
3. Four SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.

Fig. 4-3. Three SNPs within one sequence in the [dNTP1/dNTP2] format. If the SNP is marked in IUPAC code, the sequence is continued as shown in Fig. 4-4.



Please input a sequnece for SNP-RFLP analysis (the maximum load is 50)

Description: three-SNPs

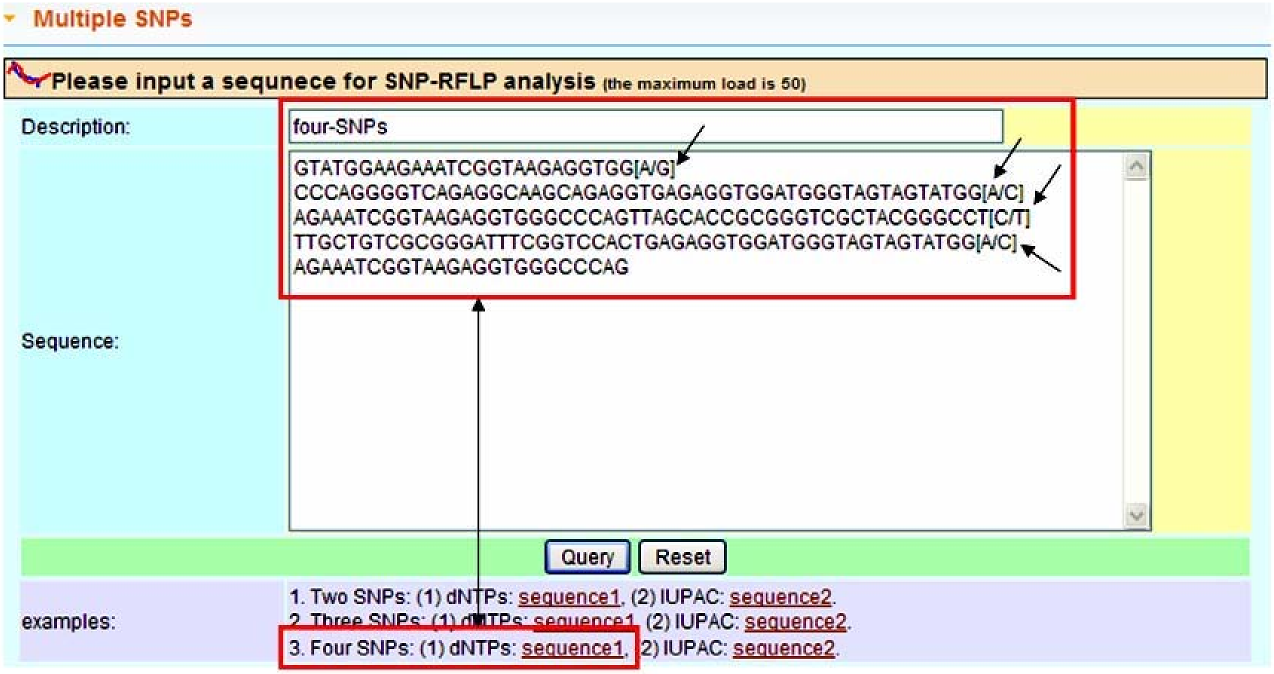
Sequence:

```
TGAGAGGTGGATGGGTAGTAGTATGGMAGAAATCGGTAAGAGGTGGGCCCAGTTAGCACCG
CGGGTCGCTACGGGCCTYTTGCTGTCGCGGGATTTCGGTCCACTGAGAGGTGGATGGGTA
GTAGTATGGMAGAAATCGGTAAGAGGTGGGCCCAG
```

examples:

1. Two SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
2. Three SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
3. Four SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.

Fig. 4-4. Three SNPs within one sequence in the IUPAC format.



Please input a sequnece for SNP-RFLP analysis (the maximum load is 50)

Description: four-SNPs

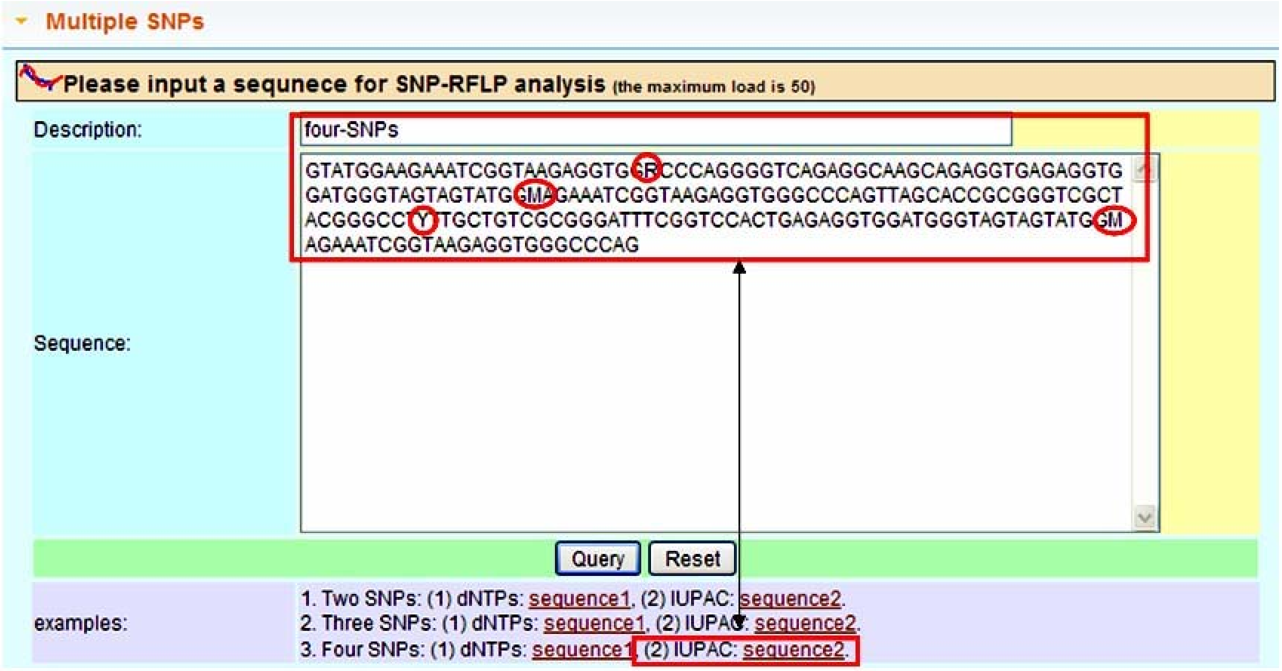
Sequence:

```
GTATGGAAGAAATCGGTAAGAGGTGG[A/G]
CCCAGGGGTCAGAGGCAAGCAGAGGTGAGAGGTGGATGGGTAGTAGTATGG[A/C]
AGAAATCGGTAAGAGGTGGGCCCAGTTAGCACCGCGGGTCGCTACGGGCCT[C/T]
TTGCTGTCGCGGGATTTCGGTCCACTGAGAGGTGGATGGGTAGTAGTATGG[A/C]
AGAAATCGGTAAGAGGTGGGCCCAG
```

examples:

1. Two SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
2. Three SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
3. Four SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.

Fig. 4-5. Four SNPs within one sequence in the [dNTP1/dNTP2] format. When the sequence is too long, the part of the sequence will list in the next line automatically. The system is able to regard them as the single sequence input. If the SNP is marked in IUPAC code, the sequence is continued as shown in Fig. 4-6.



Please input a sequnece for SNP-RFLP analysis (the maximum load is 50)

Description: four-SNPs

Sequence:

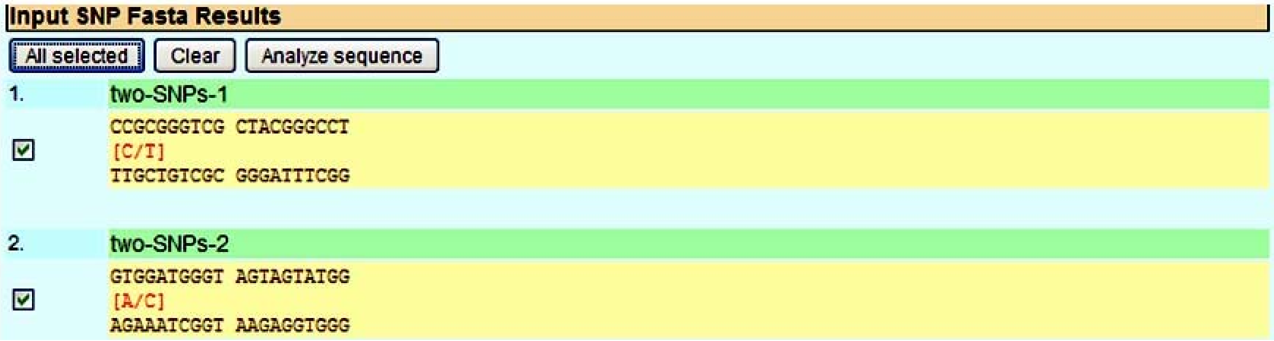
```
GTATGGAAGAAATCGGTAAGAGGTGGRCCCAGGGGTCAGAGGCAAGCAGAGGTGAGAGGTG
GATGGGTAGTAGTATGGMAGAAATCGGTAAGAGGTGGGCCCAGTTAGCACCGCGGGTCGCT
ACGGGCCTYTGCTGTCGCGGGATTTCGGTCCACTGAGAGGTGGATGGGTAGTAGTATGGM
AGAAATCGGTAAGAGGTGGGCCCAG
```

examples:
1. Two SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
2. Three SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.
3. Four SNPs: (1) dNTPs: sequence1, (2) IUPAC: sequence2.

Fig. 4-6. Four SNPs within one sequence in the IUPAC format.

After inputting multiple SNPs within one sequence, users can click the "Query" button, this system will separate multiple SNPs into specific SNP fasta sequence (Fig. 4-7). User can select the interested SNPs to to perform SNP-RFLP analysis after clicking the button "Analyze sequence". The following steps are the same as the function of **(3) SNP in fasta sequence input (page 19)** and they are shown here.

In Fig. 4-7, the "All selected" button provides convenience to check all SNPs and the "Clear" button provides cancel all SNP checked.



Input SNP Fasta Results

1. two-SNPs-1
☑
```
CCGCGGGTCG CTACGGGCCT
[C/T]
TTGCTGTCGC GGGATTTCGG
```

2. two-SNPs-2
☑
```
GTGGATGGGT AGTAGTATGG
[A/C]
AGAAATCGGT AAGAGGTGGG
```

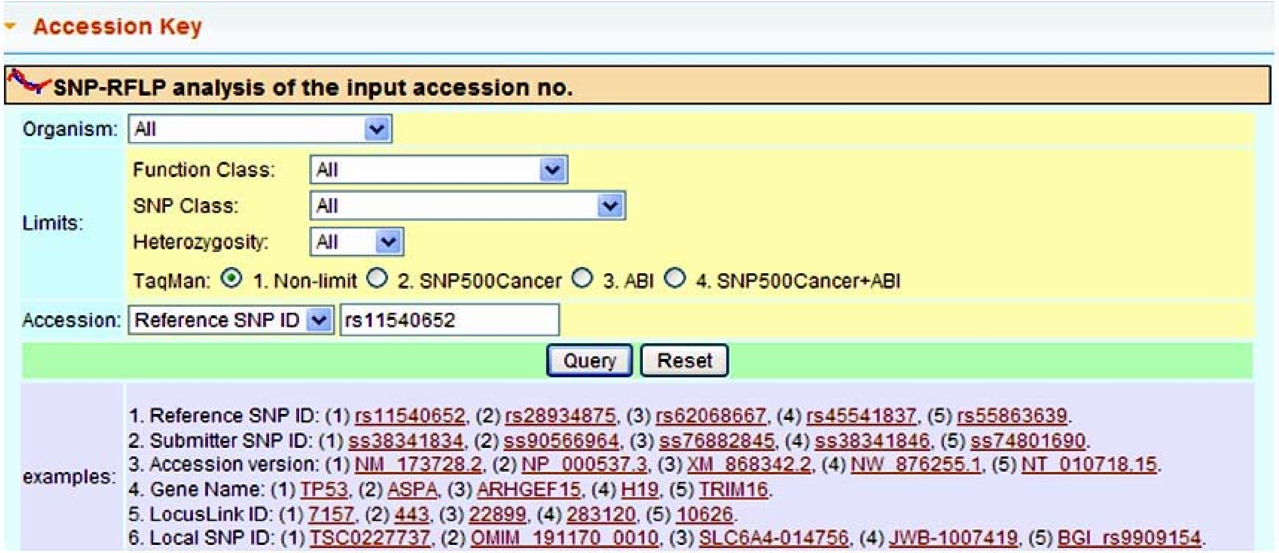
Fig. 4-7. Four SNPs within one sequence in the IUPAC format.

5. Function: GenBank accession no. for input

5.1 Overview

SNPs within the sequence information for the input accession no. of GenBank includes reference SNP ID (rs#), submitter SNP ID (ss#), accession version, HUGO gene name, local link ID (gene ID), and Local SNP ID. The classification of dbSNP in NCBI for function class (coding nonsynonymous, reference, intron, coding synonymous, locus region, mRNA UTR, and splice site), SNP class (heterozygous, indel, mixed, multinucleotide polymorphism, named locus, no variation, and snp), and heterozygosity are selectable. Furthermore, the TaqMan for non-limit, SNP500Cancer, ABI, and SNP500Cancer/ABI are available. All the information from GenBank is retrieved online for all available species (Fig. 5-1 and Fig. 5-2).

5.2 Example: Use HUGO gene name ‘TP53’ to analyze


Accession Key
SNP-RFLP analysis of the input accession no.
Query
Reset

Organism: All

Limits:
Function Class: All
SNP Class: All
Heterozygosity: All
TaqMan: ● 1. Non-limit ○ 2. SNP500Cancer ○ 3. ABI ○ 4. SNP500Cancer+ABI

Accession: Reference SNP ID rs11540652

examples:
1. Reference SNP ID: (1) rs11540652, (2) rs28934875, (3) rs62068667, (4) rs45541837, (5) rs55863639.
2. Submitter SNP ID: (1) ss38341834, (2) ss90566964, (3) ss76882845, (4) ss38341846, (5) ss74801690.
3. Accession version: (1) NM_173728.2, (2) NP_000537.3, (3) XM_868342.2, (4) NW_876255.1, (5) NT_010718.15.
4. Gene Name: (1) TP53, (2) ASPA, (3) ARHGEF15, (4) H19, (5) TRIM16.
5. LocusLink ID: (1) 7157, (2) 443, (3) 22899, (4) 283120, (5) 10626.
6. Local SNP ID: (1) TSC0227737, (2) OMIM_191170_0010, (3) SLC6A4-014756, (4) JWB-1007419, (5) BGI_rs9909154.

Fig. 5-1. This input options of GenBank Accession.

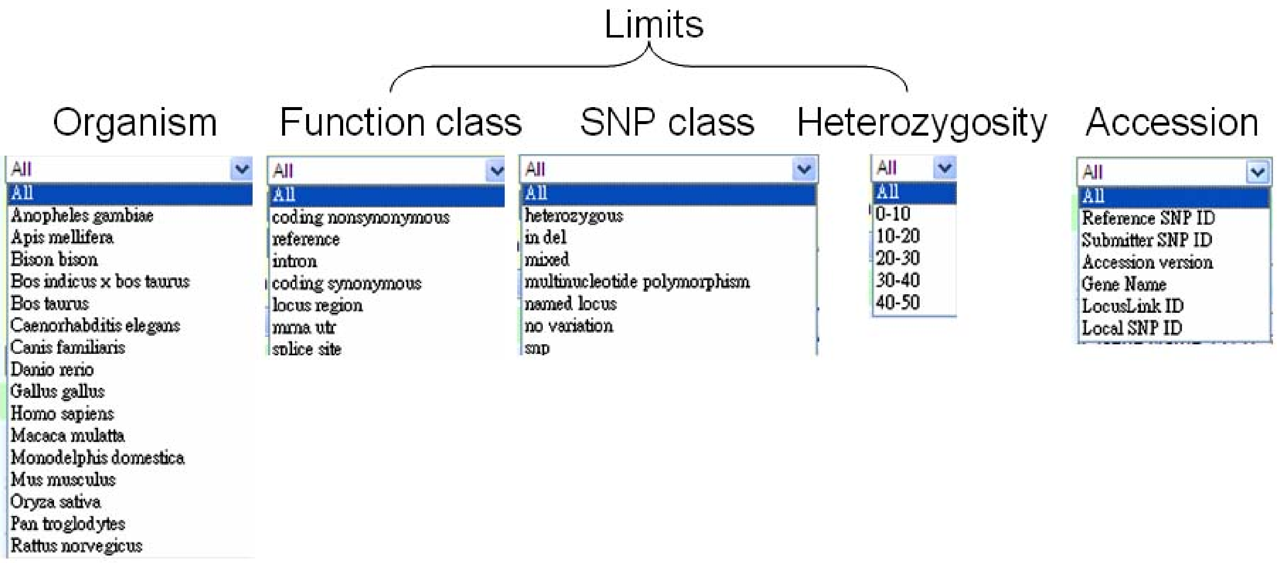
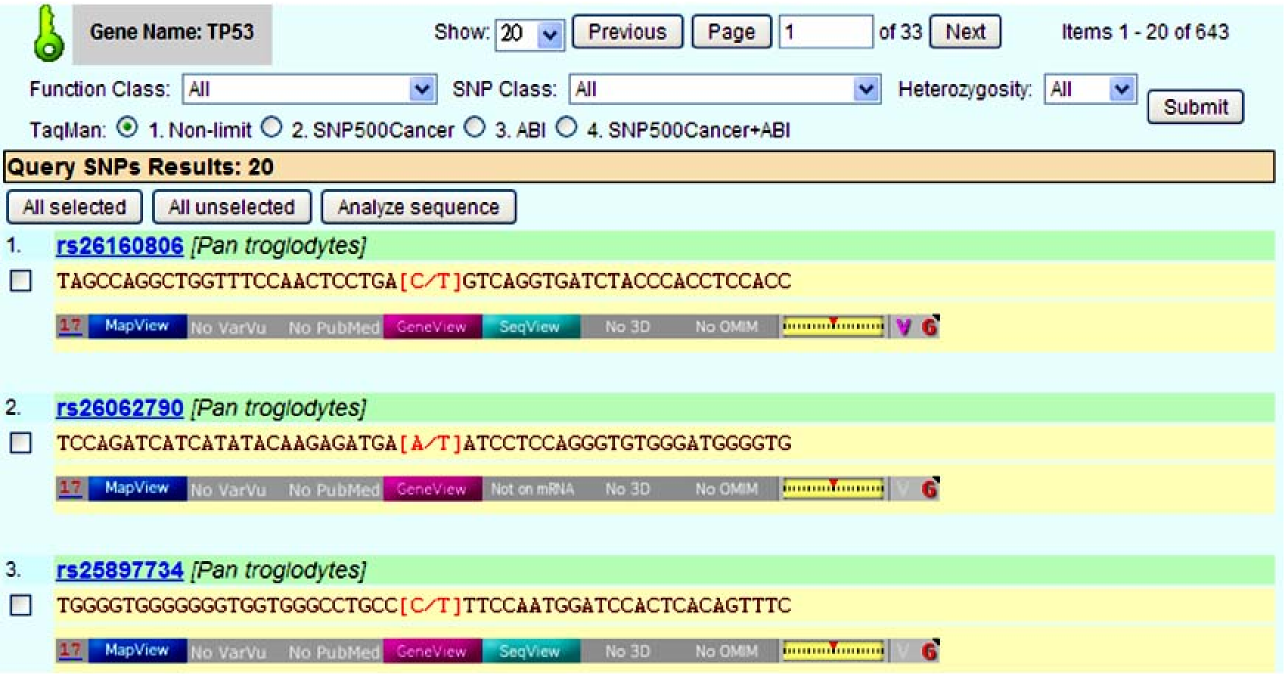

Limits
Organism
Function class
SNP class
Heterozygosity
Accession
All
All
Anopheles gambiae
Apis mellifera
Bison bison
Bos indicus x bos taurus
Bos taurus
Caenorhabditis elegans
Canis familiaris
Danio rerio
Gallus gallus
Homo sapiens
Macaca mulatta
Monodelphis domestica
Mus musculus
Oryza sativa
Pan troglodytes
Rattus norvegicus
All
All
coding nonsynonymous
reference
intron
coding synonymous
locus region
mrna utr
splice site
All
All
heterozygous
in del
mixed
multinucleotide polymorphism
named locus
no variation
snp
All
All
0-10
10-20
20-30
30-40
40-50
All
All
Reference SNP ID
Submitter SNP ID
Accession version
Gene Name
LocusLink ID
Local SNP ID

Fig. 5-2.The options for organism, limits and accession in this function of GenBank Accession.

After inputting accession no. (for example, TP53), users can click the button “Query” and it will be redirect to a waiting page (Fig. 2-12). Waiting for a moment, the results will be shown as Fig. 5-3 including: the key for input (here is Gene Name: TP53), the pages options, items for SNPs, function class, SNP class, Heterozygosity, TaqMan, and SNPs information which the same as **(2) SNP ID input** (Fig. 2-13) (page 11). The following operation (not shown) is similar as described above.



Gene Name: TP53

Show: 20 Page 1 of 33 Items 1 - 20 of 643

Function Class: All SNP Class: All Heterozygosity: All

TaqMan: ● 1. Non-limit ○ 2. SNP500Cancer ○ 3. ABI ○ 4. SNP500Cancer+ABI

Query SNPs Results: 20

1. rs26160806 *[Pan troglodytes]*
☐ TAGCCAGGCTGGTTTCCAACTCCTGA[C/T]GTCAGGTGATCTACCCACCTCCACC

2. rs26062790 *[Pan troglodytes]*
☐ TCCAGATCATCATATACAAGAGATGA[A/T]ATCCTCCAGGGTGTGGATGGGGTG

3. rs25897734 *[Pan troglodytes]*
☐ TGGGGTGGGGGGGTGGTGGGCCTGCC[C/T]TTCCAATGGATCCACTCACAGTTTC

Previous
Page
Next
Submit
All selected
All unselected
Analyze sequence
17
MapView
No VarVu
No PubMed
GeneView
SeqView
No 3D
No OMIM
V
G
17
MapView
No VarVu
No PubMed
GeneView
Not on mRNA
No 3D
No OMIM
G
17
MapView
No VarVu
No PubMed
GeneView
SeqView
No 3D
No OMIM
G

Fig. 5-3.The results for SNPs information query using HUGO gene name TP53 (here limit only show 20 SNPs and three SNPs is picked in the figure)

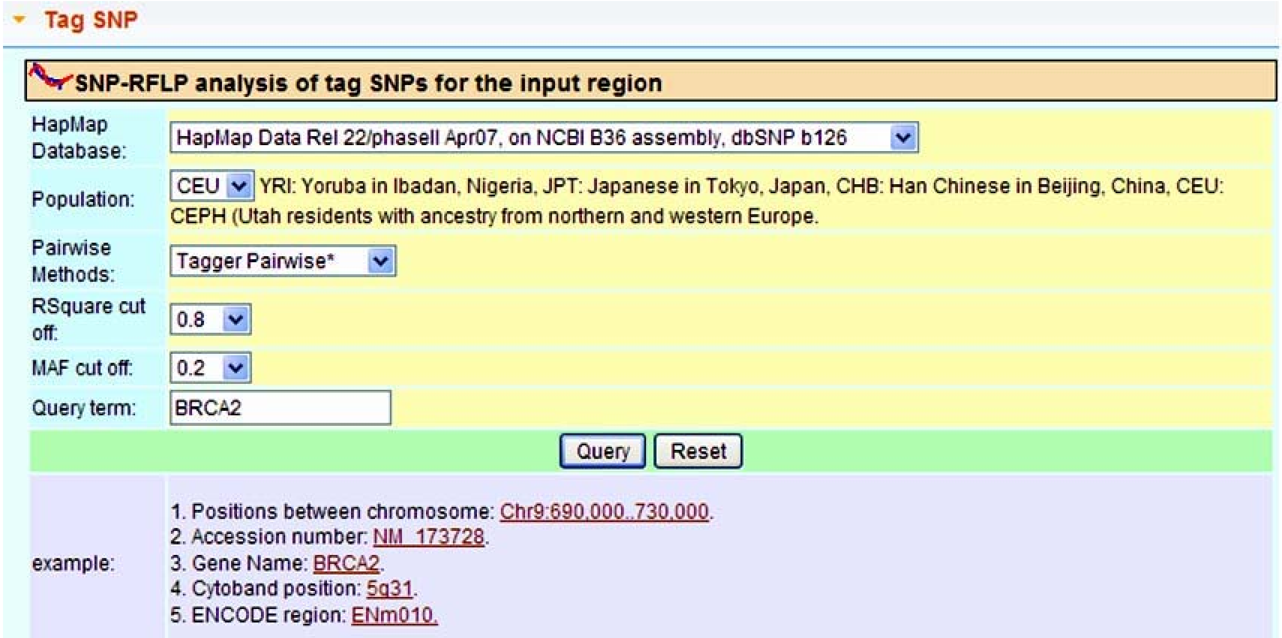
6. Function: TagSNP from HapMap

6.1 Overview

To reduce the necessary amount of SNPs for genotyping, it is believed that a subset of the SNP in a region (tagSNPs) can be chosen to represent most of the remaining SNP variants. The International HapMap Project is very helpful in providing the tagSNPs in the human genome for several races including YRI (Yoruba in Ibadan, Nigeria), JPT (Japanese in Tokyo, Japan), CHB (Han Chinese in Beijing, China), and CEU (CEPH; Utah residents with ancestry from northern and western Europe).

As shown in Fig. 6-1 and Fig. 6-2, the HapMap database versions, population, pairwise methods (tagger pairwise or tagger multimarkers), R square cut off, and MAF (minor allele frequency) cut off are adjustable. Position between chromosome, accession number, gene name, cytoband position, and ENCODE (ENCyclopedia Of DNA Elements) region can be queried. The tagSNPs information from HapMap is retrieved online and the mining function of RFLP restriction enzymes for the tagSNP is implemented. The current online linked tagSNP database is the HapMap Data Rel 23a/phaseII Mar08, on the NCBI B36 assembly, dbSNP b126. When users input query term (for example, Gene Name: ‘BRCA2’) and click the button “Query”, this system will query tagSNP from HapMap and sent a waiting page (Fig. 2-12). Then, the results for HapMap database versions, population, pairwise methods, R square cut off, and MAF cut off, and tagSNP sequence information will be shown (Fig. 6-3). The following operation is similar as described above.

6.2 Example: Use HUGO gene name ‘BRCA2’ to analyze


Tag SNP

SNP-RFLP analysis of tag SNPs for the input region

HapMap Database: HapMap Data Rel 22/phaseII Apr07, on NCBI B36 assembly, dbSNP b126

Population: CEU YRI: Yoruba in Ibadan, Nigeria, JPT: Japanese in Tokyo, Japan, CHB: Han Chinese in Beijing, China, CEU: CEPH (Utah residents with ancestry from northern and western Europe.

Pairwise Methods: Tagger Pairwise*

RSquare cut off: 0.8

MAF cut off: 0.2

Query term: BRCA2

example:
1. Positions between chromosome: Chr9:690,000..730,000.
2. Accession number: NM_173728.
3. Gene Name: BRCA2.
4. Cytoband position: 5q31.
5. ENCODE region: ENm010.

Fig. 6-1. The results for SNPs information query using HUGO gene name ‘BRCA2’ in the setting of the HapMap database of “HamMap Data Rel 22/phaseII Apr07, on NCBI B36 assembly, dbSNP b126 and other parameters as shown in this figure. The result is shown in Fig. 6-3.

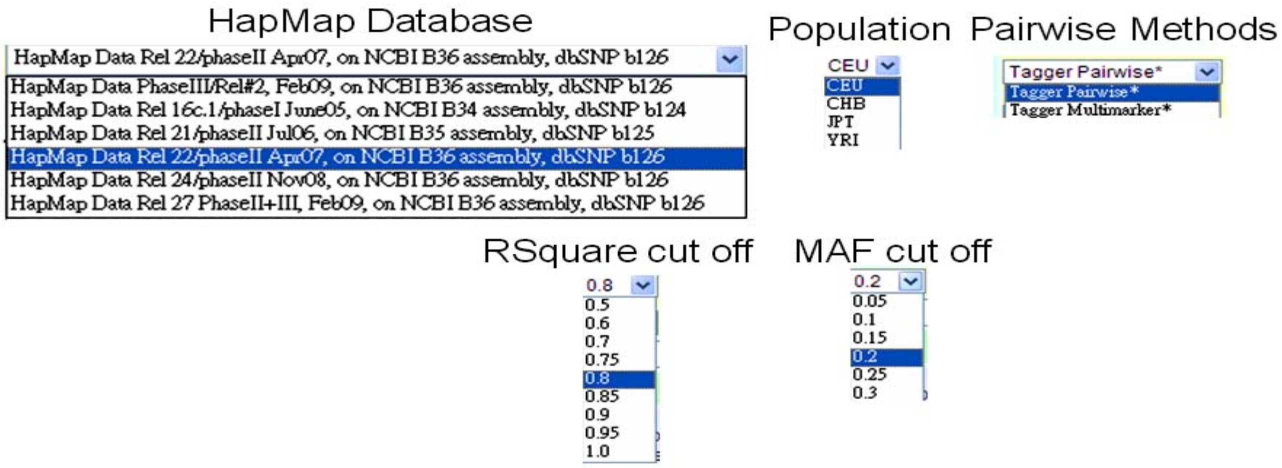
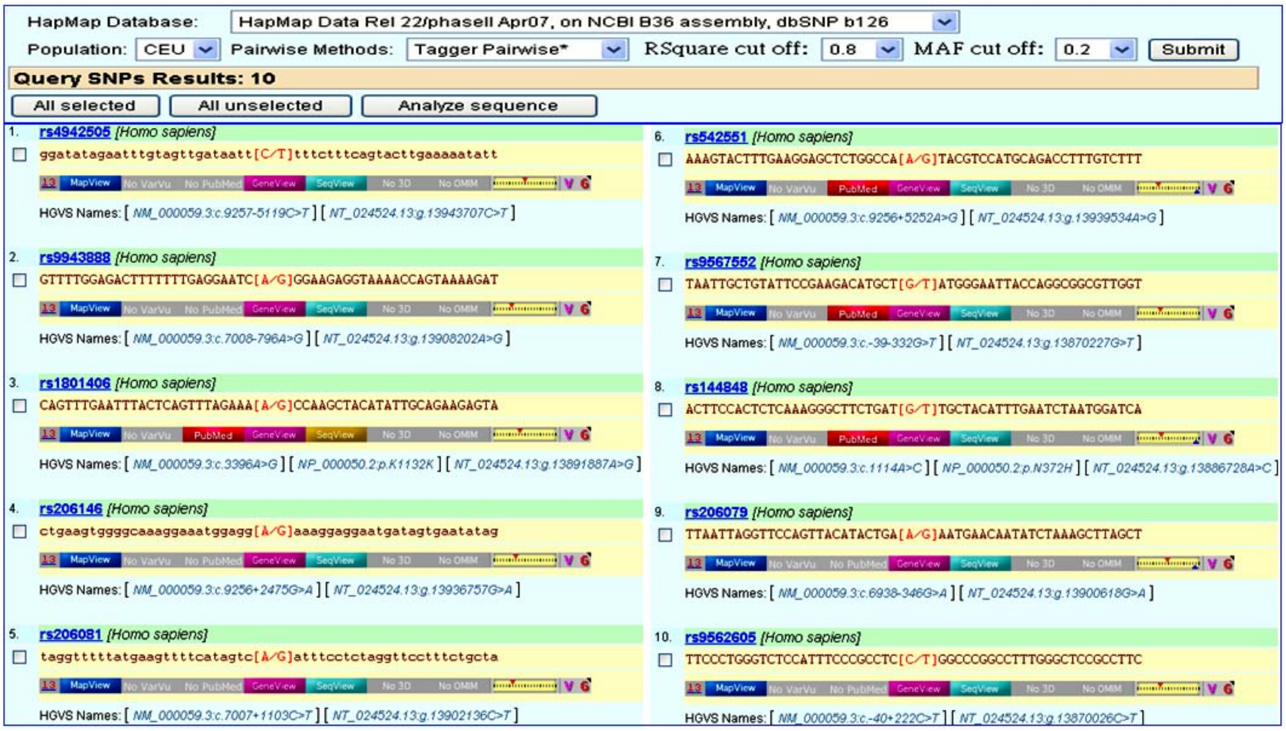

HapMap Database
HapMap Data Rel 22/phaseII Apr07, on NCBI B36 assembly, dbSNP b126
HapMap Data PhaseIII/Rel#2, Feb09, on NCBI B36 assembly, dbSNP b126
HapMap Data Rel 16c.1/phaseI June05, on NCBI B34 assembly, dbSNP b124
HapMap Data Rel 21/phaseII Jul06, on NCBI B35 assembly, dbSNP b125
HapMap Data Rel 22/phaseII Apr07, on NCBI B36 assembly, dbSNP b126
HapMap Data Rel 24/phaseII Nov08, on NCBI B36 assembly, dbSNP b126
HapMap Data Rel 27 PhaseII+III, Feb09, on NCBI B36 assembly, dbSNP b126
Population
CEU
CEU
CHB
JPT
YRI
Pairwise Methods
Tagger Pairwise*
Tagger Pairwise*
Tagger Multimarker*
RSquare cut off
0.8
0.5
0.6
0.7
0.75
0.8
0.85
0.9
0.95
1.0
MAF cut off
0.2
0.05
0.1
0.15
0.2
0.25
0.3

Fig. 6-2. The results for SNPs information query using HUGO gene name 'BRCA2' (here limit only show 20 SNPs and three SNPs is picked in the figure).


All selected
All unselected
Analyze sequence
Submit

HapMap Database: HapMap Data Rel 22/phaseII Apr07, on NCBI B36 assembly, dbSNP b126
Population: CEU Pairwise Methods: Tagger Pairwise* RSquare cut off: 0.8 MAF cut off: 0.2

Query SNPs Results: 10

1. rs4942505 *[Homo sapiens]*
☐ ggatatagaatttgtagttgataatt[C/T]tttctttcagtacttgaaaaatatt
HGVS Names: [*NM_000059.3:c.9257-5119C>T*] [*NT_024524.13:g.13943707C>T*]

2. rs9943888 *[Homo sapiens]*
☐ GTTTTGGAGACTTTTTTTGAGGAATC[A/G]GGAAGAGGTAAAACCAGTAAAAGAT
HGVS Names: [*NM_000059.3:c.7008-796A>G*] [*NT_024524.13:g.13908202A>G*]

3. rs1801406 *[Homo sapiens]*
☐ CAGTTTGAATTTACTCAGTTTAGAAA[A/G]CCAAGCTACATATTGCAGAAGAGTA
HGVS Names: [*NM_000059.3:c.3396A>G*] [*NP_000050.2:p.K1132K*] [*NT_024524.13:g.13891887A>G*]

4. rs206146 *[Homo sapiens]*
☐ ctgaagtggggcaaaggaaatggagg[A/G]aaaggaggaatgatagtgaatatag
HGVS Names: [*NM_000059.3:c.9256+2475G>A*] [*NT_024524.13:g.13936757G>A*]

5. rs206081 *[Homo sapiens]*
☐ taggtttttatgaagttttcatagtc[A/G]atttcctctaggttcctttctgcta
HGVS Names: [*NM_000059.3:c.7007+1103C>T*] [*NT_024524.13:g.13902136C>T*]

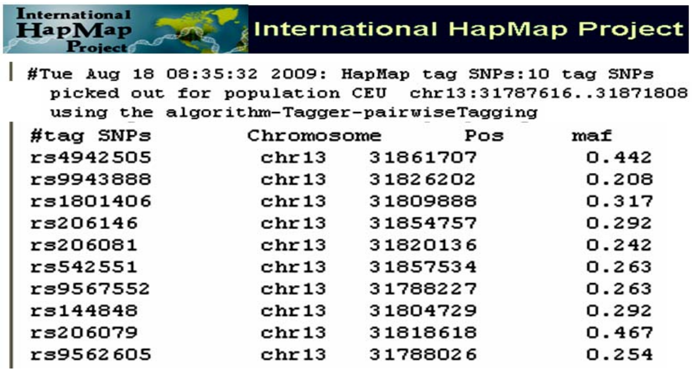
6. rs542551 *[Homo sapiens]*
☐ AAAGTACTTTGAAGGAGCTCTGGCCA[A/G]TACGTCCATGCAGACCTTTGTCTTT
HGVS Names: [*NM_000059.3:c.9256+5252A>G*] [*NT_024524.13:g.13939534A>G*]

7. rs9567552 *[Homo sapiens]*
☐ TAATTGCTGTATTCCGAAGACATGCT[G/T]ATGGGAATTACCAGGCGGCGTTGGT
HGVS Names: [*NM_000059.3:c.-39-332G>T*] [*NT_024524.13:g.13870227G>T*]

8. rs144848 *[Homo sapiens]*
☐ ACTTCCACTCTCAAAGGGCTTCTGAT[G/T]TGCTACATTTGAATCTAATGGATCA
HGVS Names: [*NM_000059.3:c.1114A>C*] [*NP_000050.2:p.N372H*] [*NT_024524.13:g.13886728A>C*]

9. rs206079 *[Homo sapiens]*
☐ TTAATTAGGTTCCAGTTACATACTGA[A/G]AATGAACAATATCTAAAGCTTAGCT
HGVS Names: [*NM_000059.3:c.6938-346G>A*] [*NT_024524.13:g.13900618G>A*]

10. rs9562605 *[Homo sapiens]*
☐ TTCCCTGGGTCTCCATTTCCCGCCTC[C/T]GGCCCGGCCTTTGGGCTCCGCCTTC
HGVS Names: [*NM_000059.3:c.-40+222C>T*] [*NT_024524.13:g.13870026C>T*]


International HapMap Project
International HapMap Project

```
#Tue Aug 18 08:35:32 2009: HapMap tag SNPs:10 tag SNPs
  picked out for population CEU  chr13:31787616..31871808
  using the algorithm-Tagger-pairwiseTagging
#tag SNPs          Chromosome        Pos        maf
rs4942505           chr13     31861707        0.442
rs9943888           chr13     31826202        0.208
rs1801406           chr13     31809888        0.317
rs206146            chr13     31854757        0.292
rs206081            chr13     31820136        0.242
rs542551            chr13     31857534        0.263
rs9567552           chr13     31788227        0.263
rs144848            chr13     31804729        0.292
rs206079            chr13     31818618        0.467
rs9562605           chr13     31788026        0.254
```

The tagSNPs provided by the HapMap are selected from different blocks. In the same block, the picked tagSNP may be different from time to time at some intervals, such as:

Interval-1 Animation *(same as top view);*

Interval-2 (partly different to Interval-1) Animation ##

##*Since the SNP-RFLPing 2 is retrieved online to the HapMap, therefore, it has the same condition as shown in the HapMap.*##

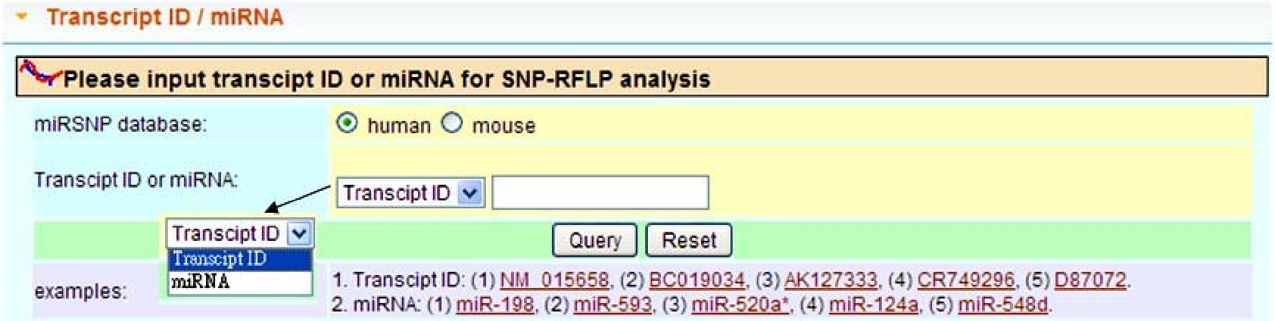
Fig. 6-3. The results for tagSNPs information query using HUGO gene name 'BRCA2' as the setting for Fig. 6-1. The ten tagSNPs retrieved in SNP-RFLPing 2 (top view; Animation) is the same as shown in HapMap (bottom view; Animation). (Please click the "Animation" box to demonstrate)

7. Function: Transcript ID/miRNA

7.1 Overview

MicroRNAs (miRNAs) are a family of small RNAs that complement to the transcripts of protein coding genes, repress translation, or decrease mRNA stability. The dysfunction of miRNA can impact many targets and, thus, affect cell physiology and cancer progression. Polymorphisms in the miRNA pathway may affect gene expression, lead to change complex phenotypes, and have the potential to be disease markers for personalized medicine. In this function, the polymorphism in microRNA Target Site (PolymiRTS) database is applied to provide transcript ID and miRNA number of human and mouse to query SNPs (Fig. 7-1).

7.2 Example: Use transcript ID and miRNA number to analyze



Transcript ID / miRNA

Please input transcipt ID or miRNA for SNP-RFLP analysis

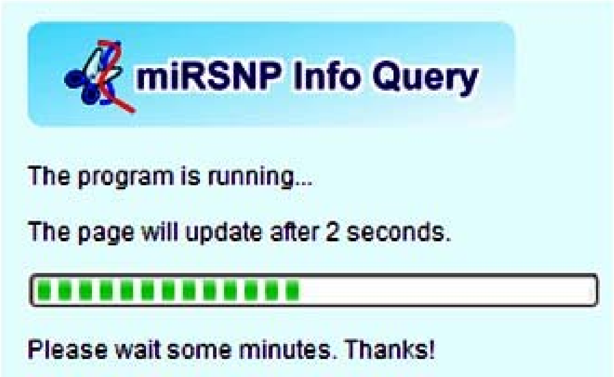
miRSNP database: ● human ○ mouse

Transcipt ID or miRNA: Transcipt ID

examples: 1. Transcipt ID: (1) NM_015658, (2) BC019034, (3) AK127333, (4) CR749296, (5) D87072.
2. miRNA: (1) miR-198, (2) miR-593, (3) miR-520a*, (4) miR-124a, (5) miR-548d.

Fig. 7-1. The interface for Transcript ID/miRNA.

When users click the button “Query”, this miRSNP information will be queried, and the waiting page will be shown (Fig. 7-2). After the query finished, the results for miRSNP will be shown. Fig. 7-3 and Fig. 7-4 show the result for query transcript ID “NM_015658” of human and the results for query transcript ID “miR-198” of human, respectively. The provided information contains transcript ID, SNP ID (rs#), Location (SNP position in transcript ID), Chr Pos (chromosome position), Ancestral Allele, Allele, miR ID (miRNA ID), Support, miR Site (miRNA site), and FuncClass (function class).

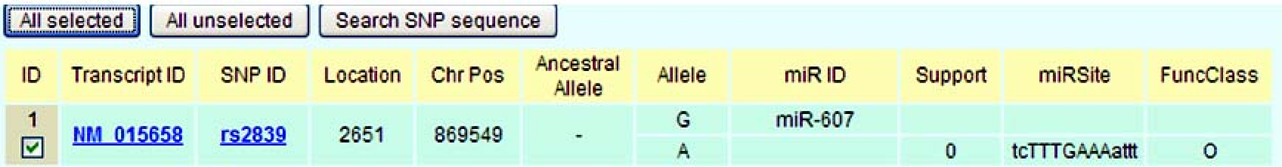


The program is running...

The page will update after 2 seconds.

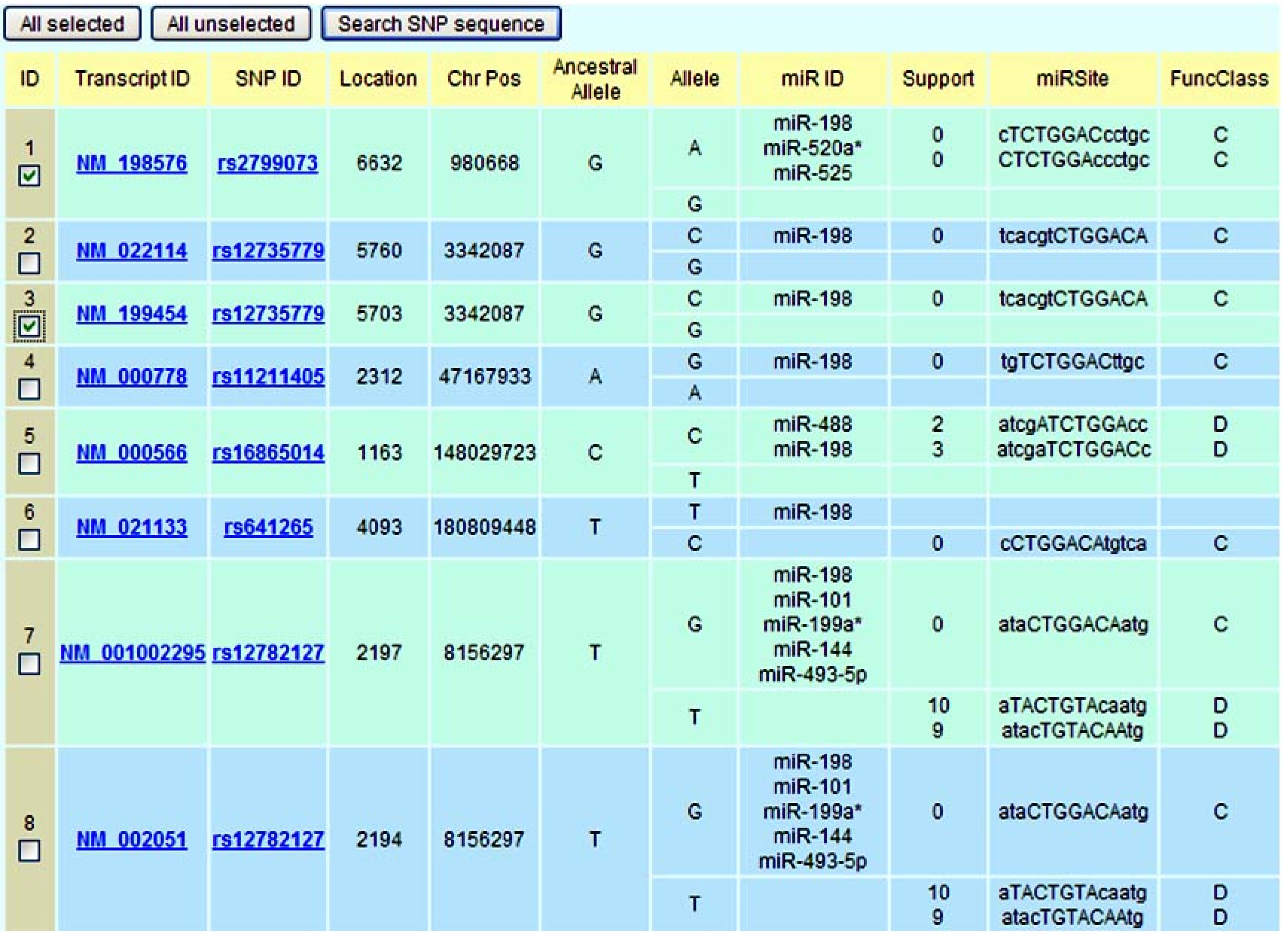
Please wait some minutes. Thanks!

Fig. 7-2. Wait for miRSNP information query.



ID	Transcript ID	SNP ID	Location	Chr Pos	Ancestral Allele	Allele	miR ID	Support	miRSite	FuncClass
1 ☑	NM_015658	rs2839	2651	869549	-	G	miR-607			
						A		0	tcTTTGAAAattt	O

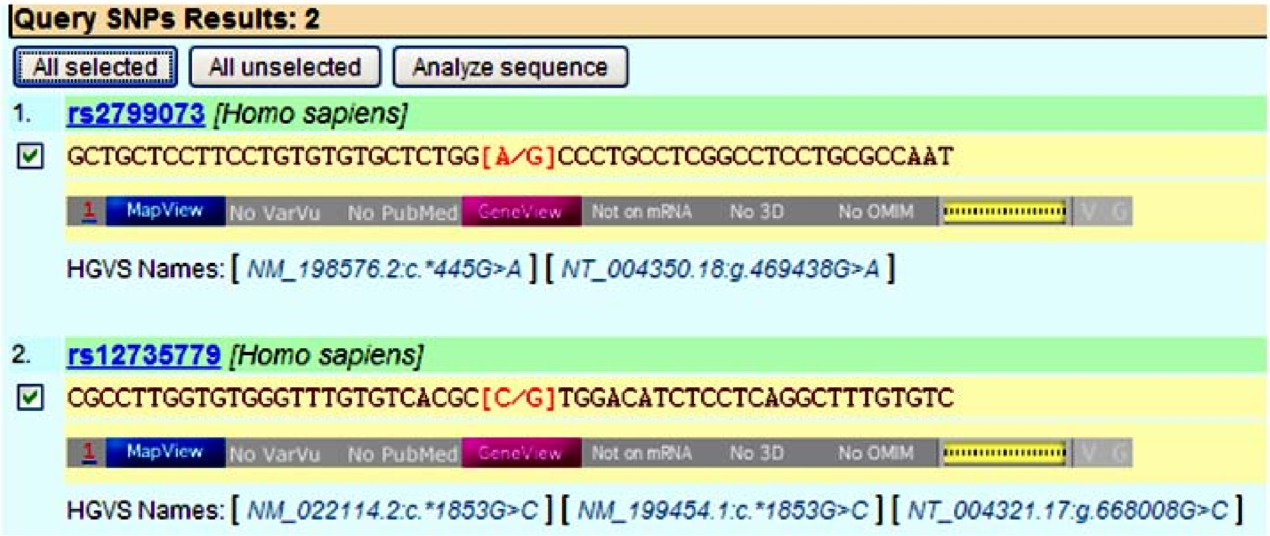
Fig. 7-3. The result for query transcript ID “NM_015658” of human.



ID	Transcript ID	SNP ID	Location	Chr Pos	Ancestral Allele	Allele	miR ID	Support	miRSite	FuncClass
1 ☑	NM_198576	rs2799073	6632	980668	G	A	miR-198 miR-520a* miR-525	0 0	cTCTGGACcctgc CTCTGGAccctgc	C C
						G				
2 ☐	NM_022114	rs12735779	5760	3342087	G	C	miR-198	0	tcacgtCTGGACA	C
						G				
3 ☑	NM_199454	rs12735779	5703	3342087	G	C	miR-198	0	tcacgtCTGGACA	C
						G				
4 ☐	NM_000778	rs11211405	2312	47167933	A	G	miR-198	0	tgTCTGGACttgc	C
						A				
5 ☐	NM_000566	rs16865014	1163	148029723	C	C	miR-488 miR-198	2 3	atcgATCTGGAcc atcgaTCTGGACc	D D
						T				
6 ☐	NM_021133	rs641265	4093	180809448	T	T	miR-198			
						C		0	cCTGGACAtgtca	C
7 ☐	NM_001002295	rs12782127	2197	8156297	T	G	miR-198 miR-101 miR-199a* miR-144 miR-493-5p	0	ataCTGGACAatg	C
						T		10 9	aTACTGTAcaatg atacTGTACAAtg	D D
8 ☐	NM_002051	rs12782127	2194	8156297	T	G	miR-198 miR-101 miR-199a* miR-144 miR-493-5p	0	ataCTGGACAatg	C
						T		10 9	aTACTGTAcaatg atacTGTACAAtg	D D

Fig. 7-4. The results for query miRNA number “miR-198” of human. (This figure show 8 records of 10 records)

In order to query the SNP sequence information to SNP-RFLP analysis, users can check the box in the front of transcript ID and click the button “Search SNP Sequence”. A waiting page for SNPs information will be displayed. Then, users can reselect SNP once again for SNP-RFLP analysis (Fig. 7-5). The following operation is similar as described above (not shown here).



Query SNPs Results: 2

1. rs2799073 [Homo sapiens]

☑ GCTGCTCCTTCCTGTGTGTGCTCTGG[A/G]CCCTGCCTCGGCCTCCTGCGCCAAT

HGVS Names: [NM_198576.2:c.*445G>A] [NT_004350.18:g.469438G>A]

2. rs12735779 [Homo sapiens]

☑ CGCCTTGGTGTGGGTTTGTGTCACGC[C/G]TGGACATCTCCTCAGGCTTTGTGTC

HGVS Names: [NM_022114.2:c.*1853G>C] [NM_199454.1:c.*1853G>C] [NT_004321.17:g.668008G>C]

All selected
All unselected
Analyze sequence
1
MapView
No VarVu
No PubMed
GeneView
Not on mRNA
No 3D
No OMIM
V
G

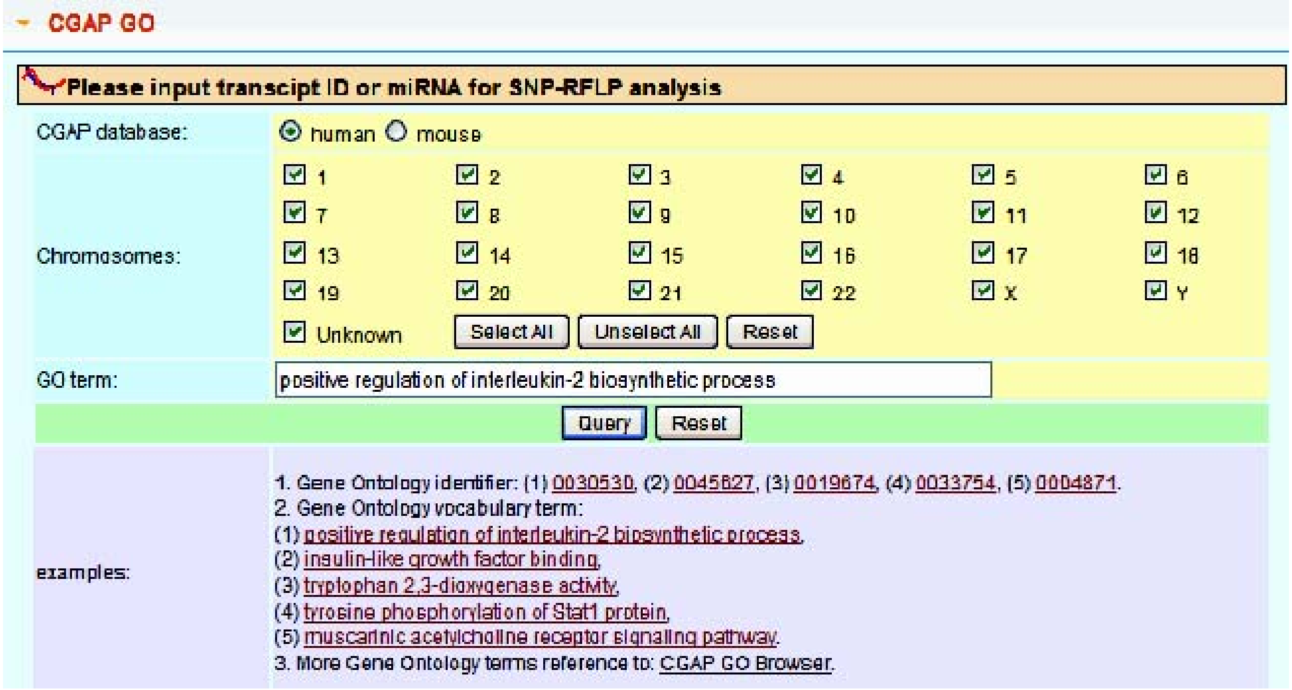
Fig. 7-5. The SNP information results for query two transcript IDs “NM_198576” and “NM_199454” of miRNA number “miR-198” of human.

8. Function: Gene Ontology-based annotation for SNPs

8.1 Overview

The Gene Ontology Browser (GO Browser; http://cgap.nci.nih.gov/Genes/GOBrowser) providing the annotation for human and mouse genes by molecular function, biological process, and cellular component. GO IDs and vocabulary terms may be input to find the genes with an interested function as well as their corresponding SNPs. In this function, CGAP database (human and mouse), Chromosomes, and GO term are available for input (Fig. 8-1). In order to get more GO terms, users can link to CGAP GO Browser (http://cgap.nci.nih.gov/Genes/GOBrowser) (Fig. 8-2).

8.2 Example: Query Gene Ontology vocabulary term



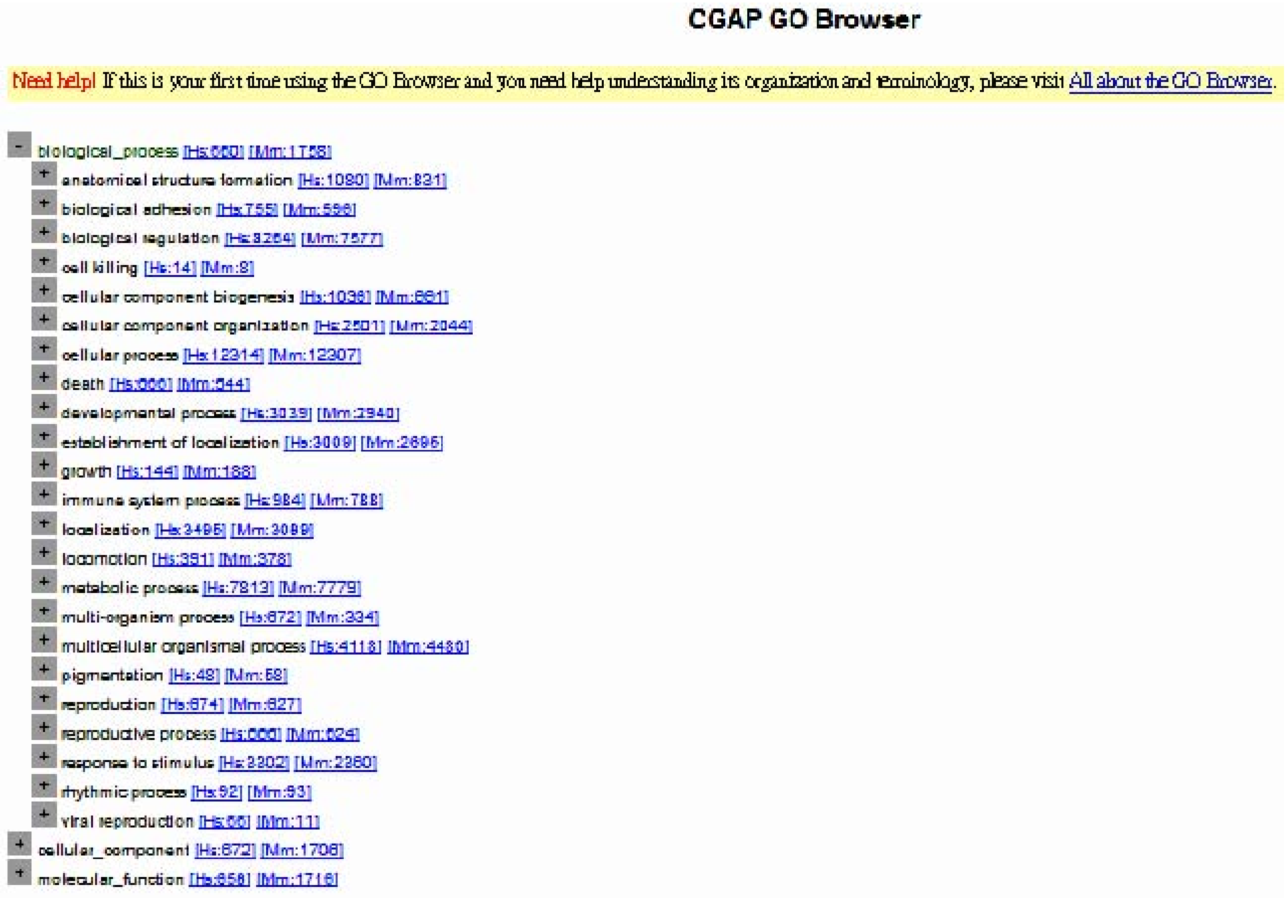
CGAP database: ● human ○ mouse

Chromosomes: ☑ 1 ☑ 2 ☑ 3 ☑ 4 ☑ 5 ☑ 6 ☑ 7 ☑ 8 ☑ 9 ☑ 10 ☑ 11 ☑ 12 ☑ 13 ☑ 14 ☑ 15 ☑ 16 ☑ 17 ☑ 18 ☑ 19 ☑ 20 ☑ 21 ☑ 22 ☑ X ☑ Y ☑ Unknown

GO term: positive regulation of interleukin-2 biosynthetic process

examples:
1. Gene Ontology identifier: (1) 0030530, (2) 0045627, (3) 0019674, (4) 0033754, (5) 0004871.
2. Gene Ontology vocabulary term:
(1) positive regulation of interleukin-2 biosynthetic process,
(2) insulin-like growth factor binding,
(3) tryptophan 2,3-dioxygenase activity,
(4) tyrosine phosphorylation of Stat1 protein,
(5) muscarinic acetylcholine receptor signaling pathway.
3. More Gene Ontology terms reference to: CGAP GO Browser.

Fig. 8-1. The interface for the input of "Gene Ontology-based annotation for SNPs".



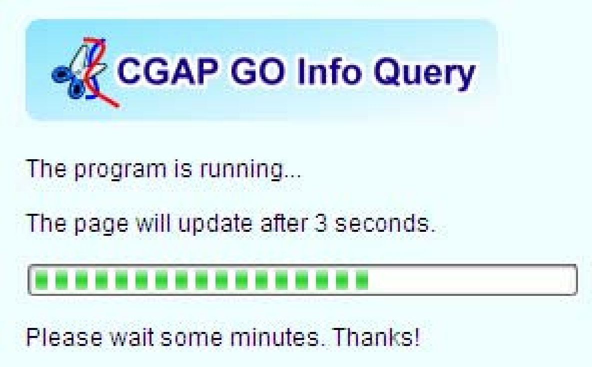
CGAP GO Browser

Need help! If this is your first time using the GO Browser and you need help understanding its organization and terminology, please visit All about the GO Browser.

- − biological_process [Hs:660] [Mm:1758]
  - + anatomical structure formation [Hs:1080] [Mm:831]
  - + biological adhesion [Hs:755] [Mm:596]
  - + biological regulation [Hs:8264] [Mm:7577]
  - + cell killing [Hs:14] [Mm:8]
  - + cellular component biogenesis [Hs:1036] [Mm:861]
  - + cellular component organization [Hs:2501] [Mm:2044]
  - + cellular process [Hs:12314] [Mm:12307]
  - + death [Hs:666] [Mm:544]
  - + developmental process [Hs:3039] [Mm:2940]
  - + establishment of localization [Hs:3009] [Mm:2696]
  - + growth [Hs:144] [Mm:188]
  - + immune system process [Hs:984] [Mm:788]
  - + localization [Hs:3495] [Mm:3099]
  - + locomotion [Hs:351] [Mm:378]
  - + metabolic process [Hs:7813] [Mm:7779]
  - + multi-organism process [Hs:672] [Mm:334]
  - + multicellular organismal process [Hs:4118] [Mm:4480]
  - + pigmentation [Hs:48] [Mm:68]
  - + reproduction [Hs:874] [Mm:827]
  - + reproductive process [Hs:666] [Mm:624]
  - + response to stimulus [Hs:3302] [Mm:2360]
  - + rhythmic process [Hs:92] [Mm:93]
  - + viral reproduction [Hs:66] [Mm:11]
- + cellular_component [Hs:672] [Mm:1706]
- + molecular_function [Hs:658] [Mm:1716]

Fig. 8-2. More GO terms can be found in CGAP GO Browser. Users can copy-and-paste the interested GO term to the window of GO term to search the SNP information for specific GO term. **Animation** (Please click the “Animation” box to demonstrate)

When users input GO term (for example, positive regulation of interleukin-2 biosynthetic process), the waiting page for querying CGAP GO information will be shown (Fig. 8-3). Waiting for a moment, the results which contain UniGene, Gene Symbol, Locus ID, Cytoband, OMIM, sequence, and SNP Information for the query GO term will be presented (Fig. 8-4). In SNPs information, two types for SNP query are available: one is NCBI SNPs, and the other is HapMap tagSNP. Through clicking the SNPs and tagSNPs images, the SNPs information will be inquired. The SNP query function is similar as described above.

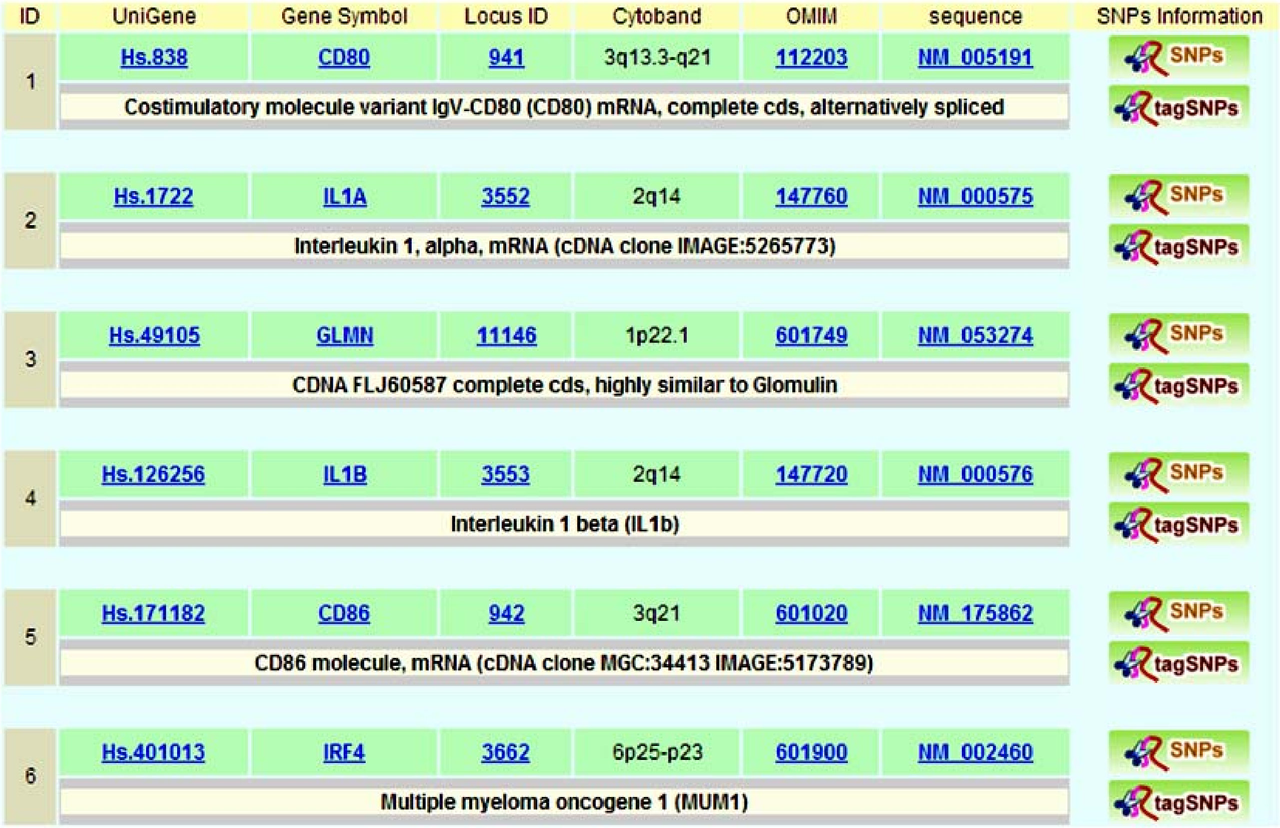


The program is running...

The page will update after 3 seconds.

Please wait some minutes. Thanks!

Fig. 8-3. Wait for CGAP GO information query.



ID	UniGene	Gene Symbol	Locus ID	Cytoband	OMIM	sequence	SNPs Information
1	Hs.838	CD80	941	3q13.3-q21	112203	NM_005191	SNPs
	Costimulatory molecule variant IgV-CD80 (CD80) mRNA, complete cds, alternatively spliced						tagSNPs
2	Hs.1722	IL1A	3552	2q14	147760	NM_000575	SNPs
	Interleukin 1, alpha, mRNA (cDNA clone IMAGE:5265773)						tagSNPs
3	Hs.49105	GLMN	11146	1p22.1	601749	NM_053274	SNPs
	CDNA FLJ60587 complete cds, highly similar to Glomulin						tagSNPs
4	Hs.126256	IL1B	3553	2q14	147720	NM_000576	SNPs
	Interleukin 1 beta (IL1b)						tagSNPs
5	Hs.171182	CD86	942	3q21	601020	NM_175862	SNPs
	CD86 molecule, mRNA (cDNA clone MGC:34413 IMAGE:5173789)						tagSNPs
6	Hs.401013	IRF4	3662	6p25-p23	601900	NM_002460	SNPs
	Multiple myeloma oncogene 1 (MUM1)						tagSNPs

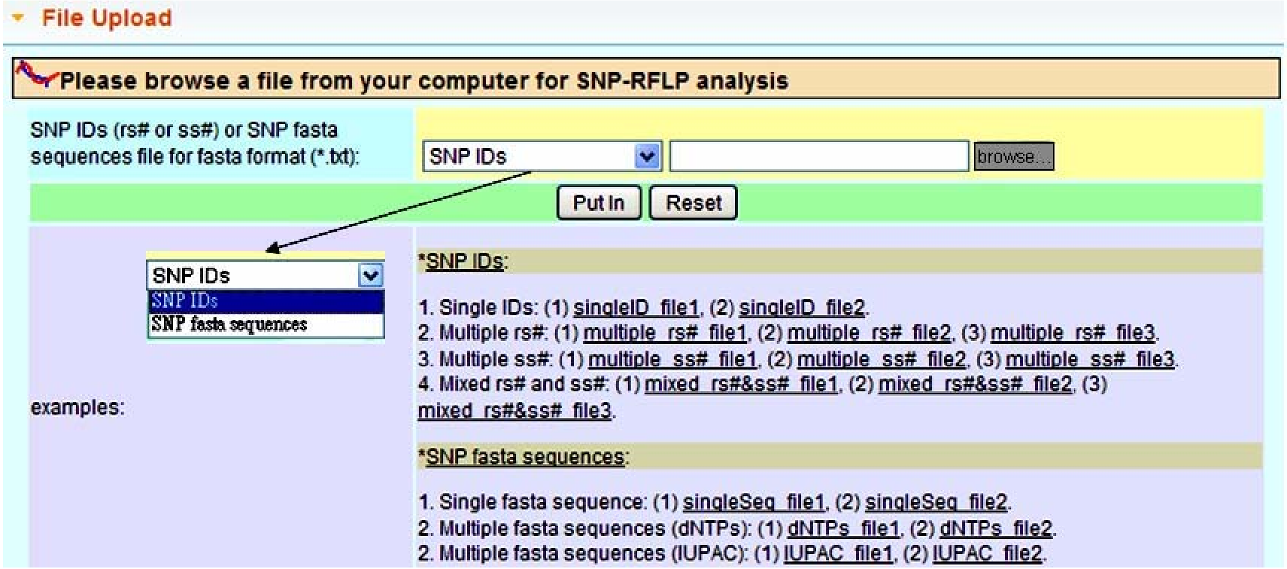
Fig. 8-4. The results for query Gene Ontology vocabulary term "positive regulation of interleukin-2 biosynthetic process".

9. Function: File upload for input

9.1 Overview

This function is similar to **(2) SNP ID input**, and **(3) SNP in fasta sequence input** functions. SNP ID (rs# and ss#) and SNP in fasta sequence format file (*.txt) are acceptable to query the SNP-RFLP information. User can select either “SNP IDs” or “SNP fasta sequences” option, and then through the button “browse…” to get an available input file in local machine (Fig. 9-1). Finally, click the button “Put In” to perform this function. For more detail formats and steps for operating, please reference to **(2) SNP ID input**, and **(3) SNP in fasta sequence input** functions.

9.2 Example: Upload interface for SNP IDs and SNP fasta sequences



Please browse a file from your computer for SNP-RFLP analysis

SNP IDs (rs# or ss#) or SNP fasta sequences file for fasta format (*.txt): SNP IDs

examples:

*SNP IDs:

1. Single IDs: (1) singleID_file1, (2) singleID_file2.
2. Multiple rs#: (1) multiple_rs#_file1, (2) multiple_rs#_file2, (3) multiple_rs#_file3.
3. Multiple ss#: (1) multiple_ss#_file1, (2) multiple_ss#_file2, (3) multiple_ss#_file3.
4. Mixed rs# and ss#: (1) mixed_rs#&ss#_file1, (2) mixed_rs#&ss#_file2, (3) mixed_rs#&ss#_file3.

*SNP fasta sequences:

1. Single fasta sequence: (1) singleSeq_file1, (2) singleSeq_file2.
2. Multiple fasta sequences (dNTPs): (1) dNTPs_file1, (2) dNTPs_file2.
2. Multiple fasta sequences (IUPAC): (1) IUPAC_file1, (2) IUPAC_file2.

Fig. 9-1. The interface for “File upload” function.

10. Appendix

10.1 Abbreviations

- **SNP:** Single nucleotide polymorphism
- **RFLP:** Restriction Fragment Length Polymorphism
- **CGAP:** Cancer Genome Anatomy Project
- **PolymiRTS:** Polymorphism in microRNA Target Site database
- **HUGO:** Human Genome Organisation
- **IUPAC:** The International Union of Pure and Applied Chemistry
- **YRI:** Yoruba in Ibadan, Nigeria
- **JPT:** Japanese in Tokyo, Japan
- **CHB:** Han Chinese in Beijing, China
- **CEU:** CEPH (Utah residents with ancestry from northern and western Europe
- **MAF:** Major Allele Frequency
- **miRNA:**
- **GO:** Gene Ontology
- **ENCODE:** ENCyclopedia Of DNA Elements

10.2 Related links

- NCBI: http://www.ncbi.nlm.nih.gov/
- REBASE: http://rebase.neb.com/rebase/rebase.html
- dbSNP: http://www.ncbi.nlm.nih.gov/SNP/
- GenBank: http://www.ncbi.nlm.nih.gov/Genbank/
- HapMap: http://www.hapmap.org/
- SNP500Cancer: http://snp500cancer.nci.nih.gov/home_1.cfm?CFID=2676524&CFTOKEN=98627459
- PolymiRTS: http://compbio.utmem.edu/miRSNP/
- CGAP GO browser: http://cgap.nci.nih.gov/Genes/GOBrowser
- ENCODE: http://genome.ucsc.edu/ENCODE/
- Prim-SNPing: http://bio.kuas.edu.tw/prim-snping/
- Seq-SNPing: http://bio.kuas.edu.tw/Seq-SNPing/
- SNP-Flankplus: http://bio.kuas.edu.tw/snp-flankplus/snpin.jsp
- LD_2SNPing: http://bio.kuas.edu.tw/LD2SNPing/
- SNP ID-info: http://bio.kuas.edu.tw/snpid-info